

Paternal and maternal lineages in the Balkans show a homogeneous landscape over linguistic barriers, except for the isolated Aromuns

E. Bosch¹, F. Calafell¹, A. González-Neira^{1,*}, C. Flaiz^{1,2}, E. Mateu¹, H.-G. Scheil³, W. Huckenbeck⁴, L. Efremovska⁵, I. Mikerezi⁶, N. Xirotiris⁷, C. Grasa⁸, H. Schmidt² and D. Comas^{1,†}

¹Unitat de Biologia Evolutiva, Universitat Pompeu Fabra, Barcelona, Spain ²Institute of Human Genetics and Anthropology, University of Ulm, Germany ³Institute of Human Genetics and Anthropology, Heinrich-Heine-University Düsseldorf, Germany ⁴Institute of Legal Medicine, Heinrich-Heine-University Düsseldorf, Germany ⁵Institute of Physiology, Medical Faculty Skopje, Republic of Macedonia ⁶Faculty of Natural Science, University Tirana, Albania ⁷Laboratory of Anthropology, Democritus University of Thrace, Komotini, Greece ⁸University Ovidius, Constanta, Romania

Summary

The Balkan Peninsula is a complex cultural mosaic comprising populations speaking languages from several branches of the Indo-European family and Altaic, as well as culturally-defined minorities such as the Aromuns who speak a Romance language. The current cultural and linguistic landscape is a palimpsest in which different peoples have contributed their cultures in a historical succession. We have sought to find any evidence of genetic stratification related to those cultural layers by typing both mtDNA and Y chromosomes, in Albanians, Romanians, Macedonians, Greeks, and five Aromun populations. We have paid special attention to the Aromuns, and sought to test genetically various hypotheses on their origins.

MtDNA and Y-chromosome haplogroup frequencies in the Balkans were found to be similar to those elsewhere in Europe. MtDNA sequences and Y-chromosome STR haplotypes revealed decreased variation in some Aromun populations. Variation within Aromun populations was the primary source of genetic differentiation. Y-chromosome haplotypes tended to be shared across Aromuns, but not across non-Aromun populations. These results point to a possible common origin of the Aromuns, with drift acting to differentiate the separate Aromun communities. The homogeneity of Balkan populations prevented testing for the origin of the Aromuns, although a significant Roman contribution can be ruled out.

Keywords: Balkan Peninsula, mitochondrial DNA, Y chromosome, genetic variation, population genetics

Introduction

The cultural and linguistic landscape in the Balkan region is complex. The Peninsula is inhabited by peoples who speak languages from several branches of the

Indo-European family, such as Greek, Albanian, Slavic (Croats, Serbians, Bulgarians and Macedonians) and Romance (Romanians and Aromuns), in addition to the Altaic Turk spoken since the Ottoman conquest of the 15th century. The Balkans have been inhabited since the Paleolithic, and the Neolithic entered Europe from Anatolia through the Balkan peninsula. The Bronze and Iron Ages were fully developed by Greeks, who controlled the eastern part of the Mediterranean until Roman rule. Romance languages spread into the Balkans with the military conquest of the Roman Empire, replacing other Indo-European languages except

*Current address: Departamento de Genética Humana, Centro Nacional de Investigaciones Oncológicas, Madrid, Spain.

†Correspondence author: David Comas, Unitat de Biologia Evolutiva, Universitat Pompeu Fabra, Doctor Aiguader 80, 08003 Barcelona, Spain. Tel: +34 93 5422844; Fax: +34 93 5422802; E-mail: david.comas@upf.edu

for the ancestors of modern Albanian and Greek. From the 6th century Slavic tribes spread over the Balkan provinces of the Byzantine Empire. It is well established that their languages replaced most Romance languages during the 9th and 10th centuries, except for Romanian that endured despite isolation from Western Romance languages, and Aromun (also known as Aromanian or Vlach), a linguistic isolate with unclear origins, scattered through the central Balkans. The Aromuns represent a small and almost unknown population that live scattered throughout the Balkans. They were previously semi-nomadic shepherds in the Balkan mountains, who settled only recently. Therefore, reliable statistical data about their size and dispersion are not available (Schmidt *et al.* 2001). Several hypotheses about the ethnogenesis of the Aromuns have been proposed (Schmidt *et al.* 2000): the Aromuns may be Latinised Greeks with some degree of Roman admixture, or the descendants of local populations living north (Dacians) or south (Thracians and Illyrians) of the Danube. However, the Aromuns represent a group defined by cultural traits such as lifestyle and language, and it remains to be proved whether they constitute a genetic unit.

Several genetic studies have provided a large amount of data on the Balkan populations, not only based on classical markers (Cavalli-Sforza *et al.* 1994) but also on autosomal STRs (Huckenbeck *et al.* 2001), Alu insertion polymorphisms (Comas *et al.* 2004), mitochondrial DNA (Calafell *et al.* 1996; Belledi *et al.* 2000; Richards *et al.* 2000; Malyarchuk *et al.* 2003) and the Y chromosome (Caglia *et al.* 1998; Rosser *et al.* 2000; Semino *et al.* 2000; Parreira *et al.* 2002; Di Giacomo *et al.* 2003; Robino *et al.* 2004). Most of these studies have focused on the impact of different human expansions in Europe, basically comparing the Paleolithic versus the Neolithic genetic contributions (Chikhi *et al.* 1998; Rosser *et al.* 2000; Semino *et al.* 2004), and/or looking at the existence of genetic structure in the extant populations of the continent (Comas *et al.* 1997; Simoni *et al.* 2000). Briefly, these studies have shown that the genetic diversity found in the Balkans fits with what is known about the European genetic landscape. However, these analyses are either European-wide or centered on one or a few Balkan populations, rather than on the Peninsula as a whole. An analysis of several Alu insertion polymorphisms in the Balkans, including Aromuns (Comas

et al. 2004), has shown that Balkan populations share a common ancestry with no major genetic barriers and a lack of correlation between genetic differentiation and language or ethnicity. Nonetheless, the analysis of Alu insertion polymorphisms has its limitations, since they do not present a well-defined phylogeographic structure of the variation. In this sense other genetic markers, such as mitochondrial DNA (mtDNA) and the Y chromosome, may provide a more detailed view of the Balkan genetic landscape.

The aim of the present study was to analyse the genetic diversity of the two uniparental markers of our genome, the mtDNA and the Y chromosome, and take advantage of their well-defined phylogenies in order to unravel the population structure of the Balkans. Moreover, for both genome regions we analysed both stable and fast-mutating polymorphisms, which respond differently to different demographic events at different time scales. Around 400 bp of mtDNA sequence and seven SNPs in the coding region were typed, as well as 19 Y chromosome STRs and 22 binary polymorphisms. MtDNA and Y-chromosome results were also compared to other published available data, to describe the overall genetic picture of the Balkans, and we attempted to test the genetic effects of the different hypotheses suggested for the ethnogenesis of the Aromuns.

Material and Methods

Subjects and Populations

Individuals from several Balkan groups were analysed for mtDNA and the Y chromosome. The sample set analyzed included: Albanians from Tirana, Greeks from Thrace, Macedonians from Skopje, Romanians from Constanta and from Ploiesti, Aromuns from Andon Poci and from Dukasi in Albania; Aromuns from Kogalniceanu in Romania, and Aromuns from Stip and from Krusevo in the Republic of Macedonia (Figure 1). Samples were taken from unrelated healthy blood donors and appropriate informed consent was obtained from all individuals participating in the study. DNA was extracted from fresh blood by either standard phenol-chloroform protocols or through the blood prep protocol recommended for extraction in the ABI PRISM™ 6700 Automated Nucleic Acid Workstation.

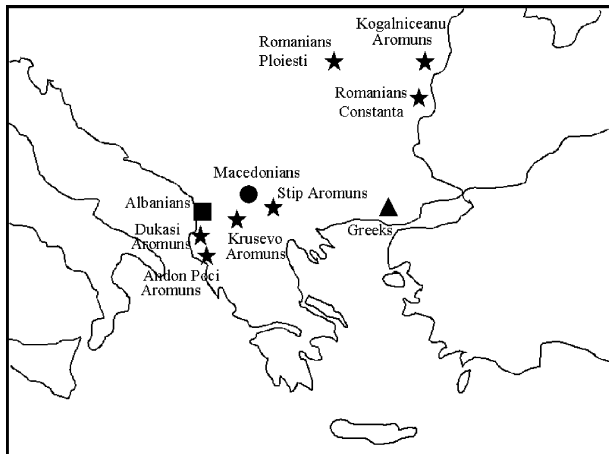


Figure 1 Geographic location of the samples analysed. Symbols represent the linguistic classification of the samples: Italic (stars), Slavic (circles), Greek (triangles), Albanian (square).

MtDNA Genotyping

MtDNA hypervariable region I was amplified using primers L15996 and H16401 (Vigilant *et al.* 1989) and the amplification products were subsequently purified with *Exo-SAP*. The sequence reaction was performed for each strand, using primers L15996 and H16401 with the ABI PRISM BigDye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems) according to supplier's recommendations. Sequences from positions 16024 to 16391 are available in Appendix I, although for most of the analyses and comparisons only the DNA stretch from positions 16024 to 16383 was considered.

Seven positions in the mtDNA coding region (7028, 10400, 10873, 11151, 11719, 12308 and 12705) were also analysed by using the SNaPshotTM ddNTP Primer Extension Kit (Applied Biosystems), which consists of a single-base primer extension that uses labelled ddNTPs to interrogate SNPs. Two mtDNA regions containing the SNPs were amplified in a multiplex reaction using primers L10373, H12744, L7008 and H7896 (see Appendix II), with the following cycling conditions: 94°C for 5 min; 35 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 30 s; and a final elongation step of 72°C for 5 min. The amplification products were purified using the QIAquickTM PCR Purification Kit (QIAGEN). The single-base primer extension was performed following supplier's recommendations, using oligonucleotides L7028X, H10400X, L10873X,

L11251X, L11719X, L12308X and L12705X in the same reaction (Appendix II). Unincorporated-labelled ddNTPs were removed by adding one unit of CIP to the primer extension products for one hour at 37°C, followed by an incubation of 15 min at 72°C to inactivate the enzyme. Products were run in an ABI PRISM377 and GeneScan Analysis Software v.3.7. was used to measure fragment sizes.

Each mtDNA molecule was assigned to one haplogroup according to the following strategy. Firstly, the combination of the seven SNPs in the coding region was taken into account to classify the mtDNA molecules into one of the eight major groups determined in the present analysis: L, M, N, R, U, HV, H, JT. Subsequently, the information yielded by the control region sequence was added, in order to refine the classification into haplogroups (Macaulay *et al.* 1999; Kivisild *et al.* 2002; Kong *et al.* 2003).

Published mtDNA sequences from several populations were used for comparison: Albanians (Belledi *et al.* 2000); Bulgarians, Greeks, Romanians from Maramures and Vrancea, Sarakatsani, Italians and Turks (Richards *et al.* 2000); Bosnians, Croatians and Serbians (Owens *et al.* 2002); and Bosnians and Slovenians (Malyarchuk *et al.* 2003).

Y-chromosome Genotyping

Y-chromosome binary polymorphisms were typed hierarchically using three different multiplex reactions. All samples were analyzed for markers M89, M172, M69, M201, M170, M9, 12f2 and M145 (multiplex I); those chromosomes assigned to clade K (M9 derived) according to the Y Chromosome Consortium (2002) were further characterised for markers M173, M45, SRY831, M207, M17 and PN25 (multiplex II), whereas those chromosomes belonging to clade DE (YAP derived branch) were further characterised for markers M96, P2, M123, M75, M78, M81, M33 and M35 (multiplex III). Amplification in multiplex was carried out in a two step PCR, using locus-specific amplification primers with a common 5'-end universal sequence at very low concentration, and adding a high concentration of universal zip code primers ZipALg1 and ZipBLg2 after 15 cycles (Appendix II). Conditions and amplification primer sequences were slightly

modified from Paracchini *et al.* (2002), except for polymorphisms 12f2 and SRY10831 whose primer sequences were modified from Blanco *et al.* (2000) and Whitfield *et al.* (1995), respectively, by adding universal code sequences at the 5' end. New amplification primers were designed for M69, M201, P2, M207, PN25 and M75 (Appendix II). PCR products were purified using SAP (Shrimp Alkaline Phosphatase) and *Exo I* (USB) to remove dNTPs and primers. The above-mentioned SNPs and the 12f2 indel were screened from the PCR generated templates by single base extension analysis, using the SNaPshot Multiplex Kit according to the manufacturer's instructions. For each set of multiplexed polymorphisms genotyping primers were designed with 4bp differing lengths (Appendix II). Unincorporated ddNTPs were removed, and the purified fragments separated and detected on a capillary electrophoresis platform as stated above.

RPS4Y₇₁₁ was typed by sequencing using primers described in Kayser *et al.* (2001) in the two chromosomes carrying ancestral states for all markers in Multiplex I.

Amplification of 19 Y STRs was performed within three multiplex reactions (MS1: DYS19, DYS388, DYS390, DYS391, DYS392 and DYS393; EBF: DYS385, DYS389 I and II, DYS460, DYS461, DYS462 and amelogenin; and CTS: DYS434, DYS435, DYS436, DYS437, DYS438 and DYS439) slightly modified from Bosch *et al.* (2002). PCR products were mixed with 400HD ROX standard and run on a ABI3100. Allele analysis and designation was carried out using haplotyped reference controls and the GeneScan Analysis Software v.3.7. Allele designation followed the nomenclature used in the Y-Chromosome Haplotype Reference Database (YHRD, <http://www.ystr.org/index.html>).

Since some of the published data on paternal lineages based on binary markers (Semino *et al.* 2000; Di Giacomo *et al.* 2003; Cinnioglu *et al.* 2004) have different phylogenetic resolution, data were homogenised in order to allow population comparisons. Data on nine Y-specific microsatellites (DYS19, DYS389 I and II, DYS390, DYS391, DYS392, DYS393, DYS385) were available for Greeks (Parreira *et al.* 2002; Robino *et al.* 2004), Albanians (Robino *et al.* 2004), Russians (Ploski *et al.* 2002), Bulgarians (Zaharova *et al.* 2001), Hungarians from Budapest (Furedi *et al.* 1999),

Italians from Rome (Caglia *et al.* 1998), Romanians (Barbarii *et al.* 2003) and Anatolian Turks (Nasidze *et al.* 2003).

Statistical Analyses

In order to detect possible genetic structure among populations, an analysis of the molecular variance (AMOVA, Excoffier *et al.* 1992) was performed using the Arlequin package ver 2.000 (Schneider, 1996; <http://anthropologie.unige.ch/arlequin/>). A correspondence analysis based on haplogroup frequencies was performed using SPSS 11.0.1 to establish the relationships between populations. This analysis provides a method for representing frequency data in Euclidian space, so that the results can be visually examined for structure (Greenacre, 1992). For data in a typical two-way contingency table, both the row variables (populations) and the column variables (haplogroups) are represented in the same space, allowing the gathering of the relationships not only from within row or column variables but also between row and column variables.

Pairwise distances for the mtDNA sequences and R_{ST} distance matrices for the Y-chromosome STRs were computed using Arlequin 2.000. The relationship between populations was also assessed through a multi-dimensional scaling (MDS) analysis, based on genetic distances among lineages and performed with the STATISTICA program. Haplogroup specific median joining networks (Bandelt *et al.* 1999) linking 19 STR haplotypes were constructed using the program Network 4.1 (<http://www.fluxus-technology.com/>). When constructing networks for each haplogroup, each STR was given a specific weight according to its variance within that haplogroup: the weight of the *i*th STR was calculated as $10V_m/V_i$, where V_m is the mean variance of all STRs and V_i is the variance of the *i*th STR. DYS385 was included in the networks as two separate loci, even if our detection procedure could not discriminate the two repeated loci; however, since the networks were built within haplogroups, it may be less likely that apparently identical DYS385 configurations were, in fact, two different genotypes. DYS385 was also included in the diversity calculations, which may have caused slight underestimates.

Results and Discussion

Mitochondrial DNA Lineages in the Balkans

The mtDNA haplogroup distribution found in the Balkans was similar to that found in other European populations (Belledi *et al.* 2000; Richards *et al.* 2000; Owens *et al.* 2002; Malyarchuk *et al.* 2003). The Balkan populations presented the characteristic European haplogroups (Table 1) with very little influence of Asian or African sequences (a maximum of 5% in Romanian Aromuns, comprising one C and one D Asian lineage). As previously described in European samples (Richards *et al.* 2000; Achilli *et al.* 2004), sequences belonging to the H haplogroup were the most prevalent in the Balkans, with frequencies around 40–50%; and with higher frequencies in Aromuns from Stip (66%). Haplogroup T1 was found at higher frequencies in Aromuns (7–14%) compared to other Balkan populations (under 7%), except for Romanians from Maramures and Vrancea (Richards *et al.* 2000) who presented similar values to those found in Aromuns. Haplogroup T1, jointly with haplogroups J and U3, have been suggested as founder Neolithic haplogroups in Europe (Richards *et al.* 2002); nonetheless, lineages belonging to the J and U3 haplogroups were found at similar levels in Aromun and non-Aromun populations.

Not only the haplogroup composition, but also haplogroup and sequence diversities and the pairwise differences, were similar in the Balkan populations compared to other European samples (Table 2). Haplogroup diversity was similar across the Balkans, except for the Aromuns from Stip who presented a reduced number of haplogroups. Nevertheless, the sequence (haplotype) diversity and the mean pairwise difference values were not significantly different between populations. These values of diversity might not properly reflect demographic processes, such as founder effects, if several distantly-related lineages are present in the founder population. In order to test whether diversity within haplogroups was substantially different among populations, the weighted mean intralinear mean pairwise difference (WIMP; Hurles *et al.* 2002), which is a measure of the diversity found within haplogroups, was calculated. Again, similar values were found, although the Aromuns from Albania (Dukasi and Andon Poci) are the populations with

the lowest WIMP values. Besides this fact, the Aromun populations presented similar mtDNA diversity values compared to the rest of the Balkan samples.

Paternal Lineages in the Balkans

The binary polymorphisms analysed split the Balkan Y chromosomes into fourteen paternal lineages (Figure 2; Table 1). Five of them were present in nearly all populations, and comprised 90% of the Y chromosomes analysed. These five haplogroups were found to characterize paternal lineages in southeastern Europe in a number of previous studies, where the geographical distribution that they encompass is also very well defined (Rosser *et al.* 2000; Semino *et al.* 2000, 2004; Wells *et al.* 2001; Barac *et al.* 2003; Cruciani *et al.* 2004).

In contrast to the most frequent haplogroups, uncommon haplogroups in the Balkans appeared in only a fraction of the populations analysed (see Figure 2). Although the most predominant paternal lineages expected in the Balkans were present in all populations analysed, their differential contributions plus the presence of some uncommon haplogroups will further characterise each population's paternal diversity patterns. When Nei's estimator of diversity was applied to haplogroup frequencies, it revealed apparent high haplogroup diversity, with only Aromuns from Albania (Dukasi and Andon Poci) presenting a slightly lower diversity (Table 3).

The analysis of 19 rapidly evolving STR loci revealed 263 STR haplotypes, among which 223 (84.79%) were found only once (see Appendix III). Haplotype diversity was ≥ 0.990 in all non-Aromun populations and Aromuns from Krusevo, in contrast to the rest of the Aromun populations (Table 3). When the WIMP measure was applied Aromuns from Albania were much less diverse, indicating a strong drift effect on their paternal lineages, as was found for their maternal lineages. Aromuns from the Stip region of Macedonia showed intermediate diversities, while other Aromun groups were indistinguishable from non-Aromuns. No Y-STR haplotypes were shared between haplogroups. However, in order to avoid other confounding effects of the phylogeny, when quantifying the haplotype differentiation among populations we also explored Y-STR haplotype variation within each population, and within the most

Table 1 MtDNA and Y-chromosome haplogroups frequencies in the Balkan populations analysed. Abbreviations: N, number of individuals; ALB, Albanians; GRE, Greeks; MAC, Macedonians; ROMC, Romanians from Constanta; ROMP, Romanians from Ploiesti; AAD, Aromuns from Dukasi in Albania; AAA, Aromuns from Andon Poci in Albania; AMK, Aromuns from Krusevo in Macedonia; AMS, Aromuns from the Stip region in Macedonia and ARO, Aromuns from Romania.

	ALB	GRE	MAC	ROMC	ROMP	AAD	AAA	AMK	AMS	ARO
MtDNA (N)	42	25	37	59	46	33	29	33	38	42
PreHV	-	-	-	0.068	-	-	-	-	-	-
HV	0.048	-	0.027	0.034	-	-	0.034	-	-	0.048
V	0.024	-	0.027	-	0.043	0.061	-	0.061	0.026	0.024
H	0.500	0.480	0.459	0.424	0.391	0.394	0.379	0.364	0.658	0.405
J*	0.048	-	0.054	0.051	0.109	0.091	-	-	-	0.048
J1	-	0.008	-	0.017	0.065	0.061	0.103	-	0.026	-
J2	-	-	-	0.017	0.022	-	-	-	-	-
T*	-	0.008	-	-	0.022	0.061	0.069	-	-	-
T1	0.024	0.004	0.054	0.034	0.043	0.121	0.069	0.121	0.079	0.143
T2	-	0.004	0.081	0.017	0.022	-	-	0.061	-	-
T3	0.024	-	0.027	-	-	-	-	-	0.026	0.024
T4	0.024	-	-	-	0.043	-	-	-	-	-
T5	-	-	-	0.017	-	-	-	0.061	-	-
K	0.048	-	0.054	0.119	0.022	0.030	0.034	0.061	0.079	0.071
U*	-	-	0.054	0.034	0.022	-	-	-	-	-
U1	0.024	-	0.027	-	-	-	-	-	0.053	-
U2	-	-	-	0.017	0.022	-	-	-	-	-
U3	-	-	-	-	0.022	-	-	0.030	0.026	-
U4	0.048	-	0.054	0.017	-	-	0.034	0.030	-	0.048
U5*	-	-	-	-	-	0.091	-	-	-	-
U5a	0.071	0.004	0.027	0.051	0.022	0.061	0.069	0.152	-	0.048
U5b	-	-	-	-	-	0.030	0.069	-	-	0.048
U7	-	-	-	-	-	-	-	-	-	-
R	-	-	-	0.034	-	-	0.034	-	-	0.024
N	0.095	0.004	0.027	0.017	0.022	-	-	-	-	-
I	-	0.004	-	-	0.022	-	0.034	-	-	-
W	0.024	0.004	0.027	-	0.022	-	0.034	0.030	-	-
X	-	0.008	-	0.034	0.065	-	0.034	0.030	0.026	0.024
Other	-	0.008	-	-	-	-	-	-	-	0.048
Y-chromosome (N)	30	41	52	31	36	39	19	43	65	42
C-RPS4Y711	-	-	-	-	-	-	-	-	-	0.048
E1-M33	-	-	-	0.032	-	-	-	-	-	-
E3b1-M78	0.233	0.171	0.212	0.097	0.139	0.179	0.158	0.186	0.185	0.071
E3b2-M81	-	-	-	-	-	-	-	0.023	-	-
E3b3-M123	-	0.024	0.019	-	0.028	-	-	-	-	-
G-M201	0.033	0.049	0.038	0.129	0.083	0.103	-	0.070	-	-
H-M69	0.067	-	-	-	0.028	-	-	-	-	-
I-M170	0.167	0.195	0.288	0.419	0.389	0.179	0.421	0.209	0.169	0.190
J(x2)-12f2	0.033	-	-	-	0.028	0.026	-	-	-	-
J2-M172	0.167	0.195	0.115	0.065	0.167	0.462	0.053	0.116	0.200	0.333
K(xP)-M9	0.033	0.024	0.038	-	-	-	-	-	-	0.095
R1*-M173×(R1a,R1b)	-	-	0.019	-	-	-	-	-	-	-
R1a1-M17	0.133	0.220	0.135	0.097	0.056	0.026	-	0.116	0.215	0.024
R1b-PN25	0.133	0.122	0.135	0.161	0.083	0.026	0.368	0.279	0.231	0.238

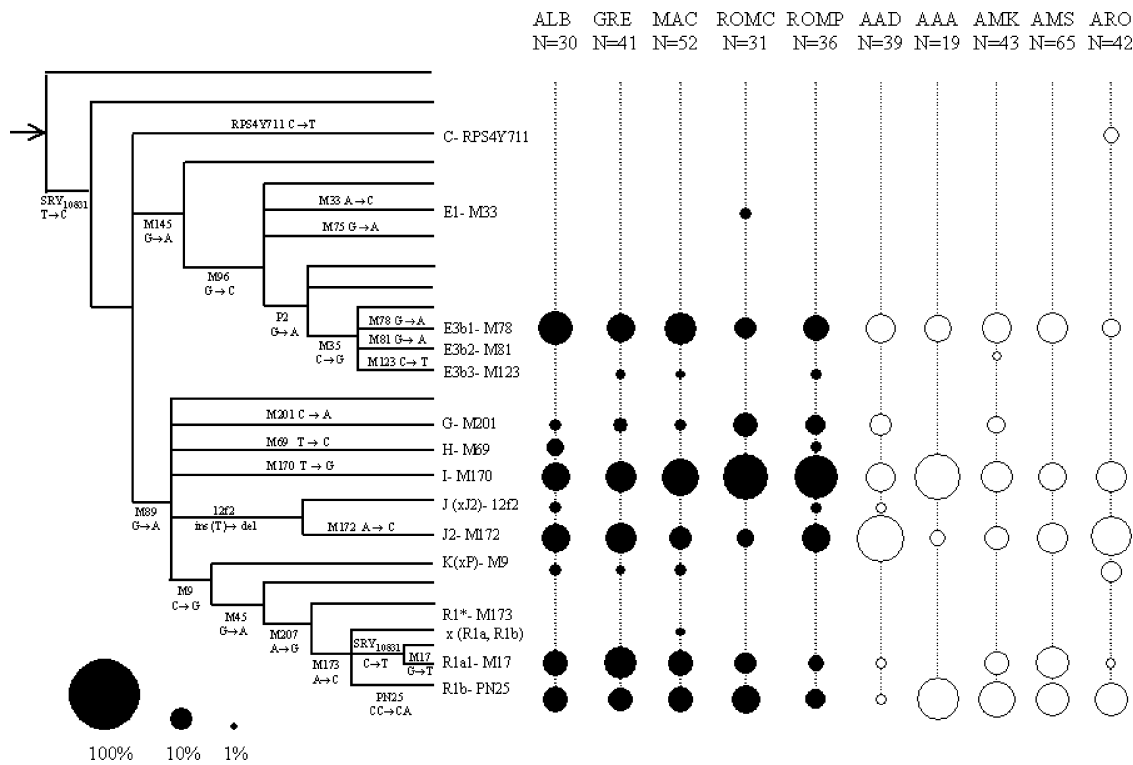
common paternal lineages, by means of median joining networks (Figure 3).

E3b1-M78 chromosomes displayed a star-like network with two Y-STR haplotypes, ht17 and ht28

(Appendix III), in its centre, separated by just one repeat difference at DYS391. DYS460 displayed a nine-repeat allele in all but six (91%) of the E3b1-M78 chromosomes analysed here. With the possible exception of

Table 2 Haplogroup and haplotype diversity values in the Balkan populations analysed for mtDNA. Abbreviations: N, number of individuals; D, diversity; MPW, mean number of pairwise differences; WIMP, the weighted intralocus mean pairwise difference.

Mitochondrial DNA	N	Haplogroups	D (haplogroups)	Sequences	D (sequences)	MPW	WIMP
Albanians	42	13	0.741 ± 0.070	30	0.968 ± 0.018	3.839 ± 1.969	1.906
Greeks	25	11	0.770 ± 0.087	20	0.963 ± 0.029	5.360 ± 2.675	2.205
Macedonians	37	14	0.784 ± 0.069	28	0.978 ± 0.014	3.934 ± 2.017	1.500
Romanians Constanta	59	17	0.803 ± 0.049	37	0.976 ± 0.009	4.259 ± 2.142	1.750
Romanians Ploiesti	46	18	0.834 ± 0.050	34	0.979 ± 0.012	4.837 ± 2.404	2.047
Aromuns							
Albania Dukasi	33	10	0.822 ± 0.056	19	0.964 ± 0.015	3.773 ± 1.951	1.495
Albania Andon Poci	29	13	0.847 ± 0.060	18	0.966 ± 0.017	4.345 ± 2.212	1.091
Macedonia Krusevo	33	11	0.837 ± 0.050	24	0.972 ± 0.016	5.447 ± 2.691	2.466
Macedonia Stip	38	9	0.563 ± 0.094	18	0.933 ± 0.022	4.366 ± 2.206	2.187
Romania Kogalniceanu	42	13	0.814 ± 0.053	33	0.988 ± 0.008	4.639 ± 2.321	2.005

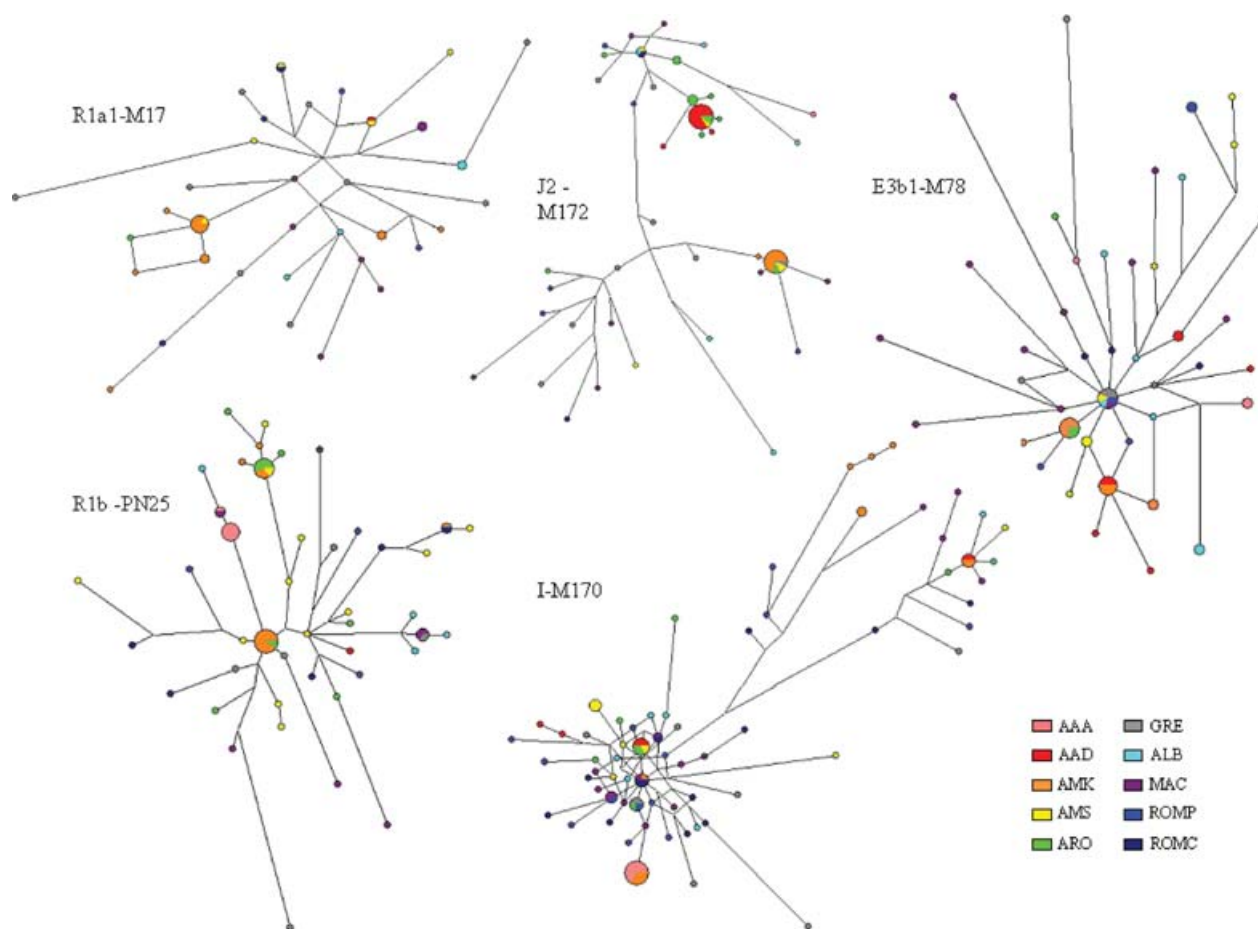
**Figure 2** Phylogeny of the Y-chromosome haplogroups found in Balkan populations. Binary markers analysed and haplogroups found are indicated on the phylogeny based on that published by the Y Chromosome Consortium (2002). To the right, representation of the haplogroup frequencies found in the Balkan populations analysed here: areas of circles are proportional to the number of chromosomes they contain. Population name abbreviations as in Table 1.

these six chromosomes, most of our E3b1-M78 chromosomes probably belong to the α cluster (Cruciani *et al.* 2004) within the E3b1-M78 lineage, characterised by the nine-repeat allele at DYS461, which has been reported as very common in the Balkans and the Aegean region. Two clusters could be distinguished in the J2-M172 network, probably reflecting the STR differentiation between any of its subclades and/or its un-

resolved paralogous J2-M172*. While the non-Aromun populations did not appear to display differences in distribution among these two clusters, the Y STR differentiation of J2-M172 chromosomes between Albanian Aromuns and the Aromun population from the Stip region in Macedonia was significant. Most of the haplotypes in the I-M170 network fell in a cluster, while the remaining were found in two main branches

Table 3 Haplogroup and haplotype diversity values in the Balkan populations analysed for the Y chromosome. Abbreviations: N, number of individuals; D, diversity; MPW, mean number of pairwise differences; WIMP, the weighted intralineage mean pairwise difference

Y chromosome	N	Haplogroups	D (haplogroups)	Haplotypes	% Individuals with private haplotypes	D (haplotypes)	MPW	WIMP
Albanians	30	9	0.876 ± 0.276	27	93.33	0.993 ± 0.011	9.901 ± 4.660	4.133
Greeks	41	8	0.849 ± 0.021	38	82.93	0.995 ± 0.007	10.220 ± 4.763	5.351
Macedonians	52	9	0.835 ± 0.026	49	84.61	0.998 ± 0.004	10.305 ± 4.779	5.386
Romanians Constanta	31	7	0.783 ± 0.058	30	80.65	0.998 ± 0.009	10.060 ± 4.726	5.866
Romanians Ploiesti	36	9	0.805 ± 0.050	35	91.67	0.998 ± 0.007	9.910 ± 4.641	5.200
Aromuns								
Albania Dukasi	39	7	0.729 ± 0.055	16	35.90	0.823 ± 0.058	8.625 ± 4.071	1.761
Albania Andon Poci	19	4	0.696 ± 0.062	6	52.63	0.743 ± 0.072	8.216 ± 3.986	0.421
Macedonia Krusevo	43	7	0.831 ± 0.025	39	74.42	0.995 ± 0.007	10.076 ± 4.696	5.496
Macedonia Stip	65	5	0.810 ± 0.010	24	29.23	0.925 ± 0.017	9.911 ± 4.593	3.324
Romania Kogalniceanu	42	7	0.798 ± 0.034	31	66.67	0.982 ± 0.010	10.943 ± 5.076	5.632

**Figure 3** Median joining networks constructed using 19-locus STR haplotypes within haplogroups E3b1-M78, J2-M172, I-M170, R1a1-M17 and R1b-PN25. Circles represent haplotypes, with areas proportional to the number of individuals they contain. Reddish and yellowish colour circles represent Aromun populations, while bluish colour circles indicate non-Aromun populations.

comprising both Aromuns and non-Aromuns. The R1a1-M17 network of haplotypes showed little structure, while R1b-PN25 was centered around a haplotype found only in Aromuns with non-Aromun haplotypes scattered all over the network. Again, the most frequent haplotypes were shared only among Aromuns.

A pattern shared by Y chromosomes in different haplogroups emerged: haplotype sharing was extensive within and between Aromun populations, but not among other Balkan populations, whose haplotypes clearly appeared scattered through the networks and separated by longer branches. This was also reflected by the fact that Aromun populations contained fewer individuals with private (population-specific) haplotypes (Table 3). Both observations could agree with a recent and common origin of Aromun paternal lineages.

Population Relationships and Genetic Structure in the Balkans

In order to visualise the relationships among the Balkan populations analysed and their surrounding neighbours, two approaches were followed: correspondence analyses based on haplogroup frequencies, and genetic distances represented in multidimensional scaling (MDS) plots.

The correspondence analysis based on mtDNA haplogroup frequencies (Figure 4a) showed a clear differentiation of the Turkish sample at one edge, characterised by haplogroups frequent in the Middle East. The rest of samples were clustered in the centre of the plot, with the exception of two Aromun samples from Dukasi in Albania and Krusevo in Macedonia, which were detached from the rest of samples. When the correspondence analysis based on Y-chromosome haplogroup frequencies was performed with the whole set of populations for comparison (Figure 4b), all the Balkan populations analysed in the present study, plus the additional Macedonians, Albanians, Italians, Greeks and the two Turkish samples clustered more or less together, separated from other Slavic populations (Croatians, Polish, Ukrainians and Czech-Slovakians) and the Hungarian sample that formed a more differentiated group.

In order to represent the genetic distances between samples, an MDS was performed with the full data from mtDNA sequences. The plot of the first two dimensions of the MDS clustered most of the populations

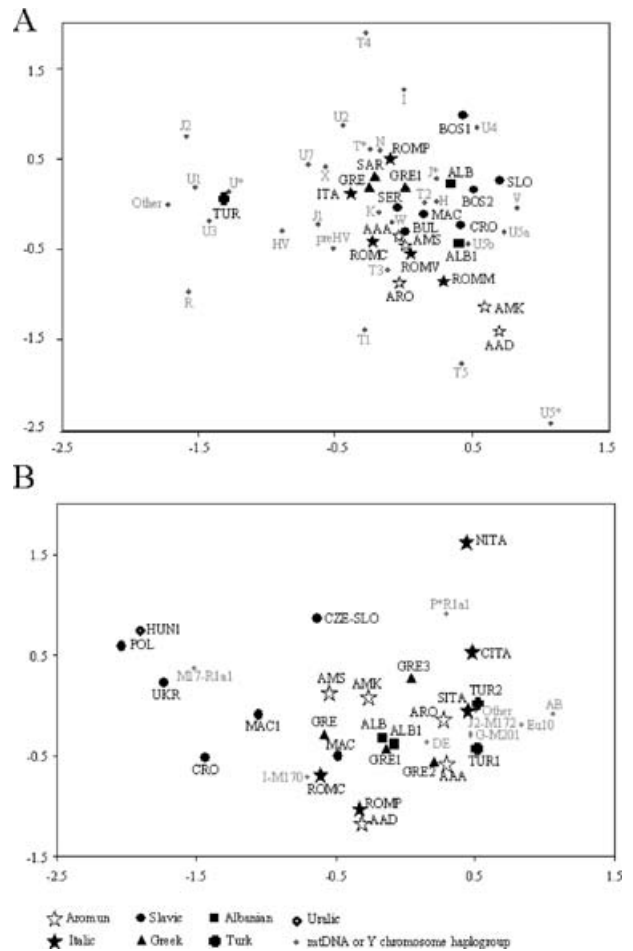


Figure 4 Plot of the first two dimensions in the correspondence analysis from mtDNA (a) and Y-chromosome (b) haplogroup frequencies. Symbols represent the linguistic classification of the samples: Italic (stars), Slavic (full circles), Greek (triangles), Albanian (squares), Uralic (open circles), Turk (crosses). Abbreviations for populations typed in this study as in Table 1. Abbreviations for populations used for comparison: A) mtDNA: ALB1, Albanians (Belledi *et al.* 2000); GRE1, Greeks (Richards *et al.* 2000); ROMM and ROMV, Romanians from Maramures and Vrancea, respectively (Richards *et al.* 2000); BOS1 (Owens *et al.* 2002) and BOS2 (Malyarchuk *et al.* 2003), Bosnians; SLO, Slovenians (Malyarchuk *et al.* 2003); CRO, Croatians (Owens *et al.* 2002); SER, Serbians (Owens *et al.* 2002); TUR, Turks (Richards *et al.* 2000); ITA, Italians (Richards *et al.* 2000); SAR, Sarakatsani (Richards *et al.* 2000). B) Y-chromosome: ALB1, Albanians; CRO, Croatians; CZE-SLO, Czechs and Slovakians; HUN1, Hungarians; MAC1, Macedonians; POL, Polish; TUR1, Turks; and UKR, Ukrainians (all from Semino *et al.* 2000); TUR2, Anatolian Turks (Cinnioglu *et al.* 2004); NITA, CITA, SITA, North, Central and South Italians respectively (Di Giacomo *et al.* 2003); GRE1, Continental Greeks (Di Giacomo *et al.* 2003) and GRE2, Cretean Greeks (Di Giacomo *et al.* 2003) and GRE3, Greeks (Semino *et al.* 2000).

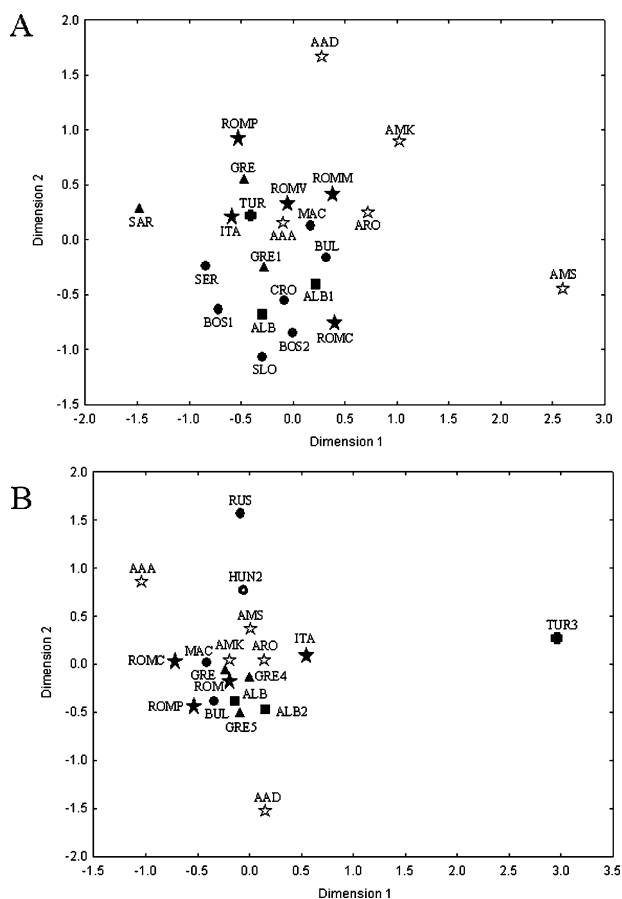


Figure 5 Multidimensional scaling analysis from distance matrices a: Pairwise distances between mtDNA sequences. b: R_{ST} distances between Y-chromosome STR haplotypes. Abbreviations as in Figure 4. Stress values were 0.132 for mtDNA and 0.096 for the Y chromosome. Other samples: RUS, Russians (Ploski *et al.* 2002); GRE4 (Parreira *et al.* 2002) and GRE5 (Robino *et al.* 2004), Greeks; ALB2, Albanians (Robino *et al.* 2004); BUL, Bulgarians (Zaharova *et al.* 2001); HUN2, Hungarians from Budapest (Furedi *et al.* 1999); ITA, Italians from Rome (Caglia *et al.* 1998); ROM, Romanians (Barbarii *et al.* 2003), and TUR3, Anatolian Turks (Nasidze *et al.* 2003).

analysed, whereas the Aromun samples, especially the Aromuns from Dukasi, Stip and Krusevo, remained in the periphery of the plot (Figure 5a). On the other hand, the MDS plot based on R_{ST} distances, constructed from 9 STRs haplotypes in the Y chromosome, showed that Turks from Anatolia were highly differentiated from the Balkan region while Albanian Aromuns (from Andon Poci and Dukasi) were distinctly separated from the remaining Balkan populations (Figure 5b). This fact was also observed in an MDS plot based on the R_{ST} distances between 19 STR haplotypes (data not shown), considering only the samples typed in the present study.

It is worth noting the central location of the Romanian Aromuns in all plots. This group was constituted two generations ago by Aromun immigrants from different Balkan origins (Schmidt *et al.* 2001), which explains their short genetic distances to other Aromun groups.

The analyses of population relationships based on stable and fast polymorphisms both in mtDNA and in the Y-chromosome showed that most of the Balkan populations form a homogeneous set, and are similar to surrounding populations. However, the faster evolving loci showed that some particular Aromun groups are differentiated from this common background. It should be noted that this is a descriptive analysis lacking any formal testing. In order to test the genetic structure in the Balkans, analyses of the molecular variance (AMOVA) were performed, considering the haplogroups or haplotypes of the mtDNA and the Y chromosome.

No significant differences were found considering mtDNA haplogroups in the AMOVA, neither considering all populations as a single group nor grouping populations according to Aromun affiliation, language or country (Table 4). When the same analysis was performed on the mtDNA sequences, significant differences were found between Aromuns and non-Aromuns, as well as significant heterogeneity found within Aromun samples. Again, the genetic sequence diversity apportioned by grouping Balkan samples by language or country was not significant. The AMOVA for the Y chromosome showed significant differences in haplogroup and 19-STR haplotype composition between populations in the Balkan region. These differences were mainly the result of the heterogeneity found in the Aromun populations, given that non-Aromun populations alone did not show significant differentiation among themselves (Table 4). No significant differences ($p > 0.05$) were found either between Aromuns and non-Aromun populations or between different groupings based on country or language affiliation.

This set of analyses, in agreement with the graphical description presented above (i.e. correspondence analysis and MDS), shows that the main source of genetic differentiation in the Balkans is due to some, but not all, Aromun groups. This pattern is more evident in the fast evolving sites.

Table 4 Analysis of Molecular Variance (AMOVA) in the Balkan populations analysed using mtDNA and Y-chromosome data. Language: all populations grouped by language. Country: populations grouped by country. HG: haplogroups, HT: haplotypes, ns not significant, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

AMOVA		Among groups	Among populations	Within populations
MtDNA (HG)	Balkans		0.51 ns	99.49
	Non-Aromun		-0.31 ns	100.31
	Aromun		1.40 ns	98.60
	Aromuns vs non-Aromun	0.08 ns	0.47 ns	99.45
	Aromuns vs Romanians	-0.09 ns	0.97 ns	99.11
	Aromuns vs Greeks	-1.26 ns	1.42 ns	99.84
	Language	-0.90 ns	0.96 ns	99.94
	Country	-0.30 ns	0.75 ns	99.55
MtDNA (HT)	Balkans		1.16 ***	98.84
	Non-Aromun		0.63 ns	99.37
	Aromun		1.17 *	98.83
	Aromuns vs non-Aromun	0.49 *	0.88 ***	98.63
	Aromuns vs Romanians	0.16 ns	1.25 *	98.60
	Aromuns vs Greeks	0.48 ns	1.07 *	98.45
	Language	-0.50 ns	1.41 *	99.09
	Country	-0.17 ns	1.29 ns	98.88
Y-chr. (HG)	Balkans		2.95***	97.05
	Non-Aromun		0.26 ns	99.74
	Aromun		4.78***	95.22
	Aromuns vs non-Aromun	0.65 ns	2.58***	96.77
	Aromuns vs Romanians	2.05 ns	3.69***	94.26
	Aromuns vs Greeks	-2.74 ns	4.82***	97.92
	Language	-2.96 ns	4.61**	98.35
	Country	-0.51 ns	3.35**	97.15
Y-chr. (HT) 19 STRs	Balkans		5.52***	94.48
	Non-Aromun		0.23 ns	99.77
	Aromun		9.54***	90.46
	Aromuns vs non-Aromun	0.34 ns	5.31***	94.34
	Aromuns vs Romanians	-0.36 ns	8.45***	91.91
	Aromuns vs Greeks	-6.06 ns	10.21***	95.86
	Language	-5.14 ns	8.46**	96.67
	Country	-0.23 ns	5.70***	94.53

Drift in the Aromuns

All the Balkan populations analysed here were genetically homogeneous with the exception of some Aromun samples. This was particularly evident with the Y chromosome, as both haplogroup and 19 STR haplotype based data showed significant differences ($p < 0.001$) among the Aromun groups. Therefore, it seems that the Aromun populations do not constitute a homogenous group separated from the rest of the Balkan populations, but that they present relative heterogeneity, especially for paternal lineage composition, between themselves. The non-significance of the paternal differentiation of Aromun populations versus non-Aromuns is probably due to this high Aromun heterogeneity, meaning that most genetic distances between any Aromun group pair are

greater than those between any non-Aromun population pair. In spite of their possible historical common origin, the geographical isolation between the Aromun populations analysed, plus the cultural isolation from their neighbours, may have favoured the action of genetic drift on their paternal lineage composition even at the level of binary markers. The reduction in both haplogroup and haplotype internal diversity values in some Aromun populations also agrees with the action of drift. Moreover, each Y chromosome haplogroup can be taken as an independent view of the evolutionary process, and does not have to display exactly the same pattern given the randomness of the evolutionary process. However, the repeated pattern in all the paternal lineages found in shared haplotypes between

Aromun populations, and the low fraction of individuals with private haplotypes among the Aromuns, provides further evidence of the effect of genetic drift in these populations.

Although mtDNA haplogroup composition was not significantly different among the Aromun groups, we observed certain haplogroup and haplotype diversity reduction for those Aromuns from Stip in Macedonia, plus significant differences ($p < 0.05$) in mtDNA haplotype composition among the Aromun populations (Table 4). Particular Aromun groups, such as those from Andon Poci and Dukasi in Albania showed lower haplogroup and haplotype diversities with the Y chromosome, while their mtDNA diversities were similar to those of other Aromun groups and Balkan populations. Ethnological observations for those localities evidenced a pattern of patrilocality, with an inflow of Aromun women from other regions and villages. This process would replenish mtDNA diversities while maintaining high levels of Y-chromosome drift.

Origins of the Aromuns

Several hypotheses have been proposed for the origin of the Aromuns (Schmidt, 2000; Comas *et al.* 2004): i) Aromuns are Latinised Greeks; ii) Aromuns (and present Romanians) are descendants of Dacians (who lived north of the Danube), or iii) Aromuns are descendants of Thracians (who lived south of the Danube). Moreover, all three scenarios could have introduced different degrees of Roman admixture. All three scenarios might have produced different genetic outcomes, with being Aromuns being genetically closer to Greeks, Romanians, or other populations from the south of the Balkans, respectively. However, the possibility of testing these hypotheses depends on these populations being genetically differentiated. While Y chromosome variation has been previously reported as being very well structured geographically (Rosser *et al.* 2000; Semino *et al.* 2000), mtDNA lineages are believed to be much more homogeneous between geographically close populations (Simoni *et al.* 2000). Besides, our ability to discern the possible genetic origin of the Aromuns also depends on the action of drift, as this may have erased such traces or their origins.

As shown in the AMOVAs (Table 4), Aromuns as a group were not significantly different either from Romanians or from Greeks, considering both haplotypes and haplogroups for both uniparental markers. The comparison of genetic distances among the different Aromun groups and their surrounding neighbours, when using mtDNA sequences, revealed great genetic homogeneity in the Balkan peninsula and in particular no significant differences ($P < 0.05$) among Greeks, Italians, Romanians and Bulgarians or Macedonians (the last two being taken as proxies for other southern Balkan populations). This result implies that the power to test the proposed hypotheses is not sufficient, at least for the maternal lineages. The Balkan region also appears relatively homogenous when using both Y chromosome STRs or haplogroup frequency data. The comparison of R_{ST} distances when using Y chromosome 9-STR haplotypes displayed no significant differences ($P > 0.05$) among Greeks, Romanians, Macedonians or Bulgarians implying that we cannot distinguish between the second and third hypotheses. However, we did have enough statistical power to test a Roman (Italian) paternal contribution in the Aromun populations. Therefore, we could only distinguish the first hypothesis if “Latinised Greeks” implied a major Roman (Italian) contribution to the Aromuns.

Y STR based R_{ST} distances between most Aromun groups and Italians were shorter than those between Italians and Greeks, although all Aromuns groups are much closer to Greeks than to Italians. Visual inspection revealed that the only Y chromosome lineage that had frequencies in the Aromuns that were closer to those in Italians than in the rest of the Balkans was $P^*(\times R1a)$. These frequencies were between 0.057–0.279 in the Aromuns, 0.083–0.161 in the rest of the Balkans and 0.240–0.630 in Italy. This situation could indicate paternal gene flow mediated by the Romans, as foreseen by the first hypothesis. However, when the YHRD database was queried for the nine-locus “minimal haplotype” (DYS19, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS385I/II) corresponding to the $P^*(\times R1a)$ Y chromosomes found in the Aromuns, it showed few instances of matches in Italians but not in the Balkans. On the contrary, nine $P^*(\times R1a)$ Aromun minimal haplotypes, belonging to a total 18 individuals, had no matches in the YHRD. This

indicates that the increased frequencies of P* (× R1a) Y chromosomes in the Aromuns seem to be due to drift rather than to external gene flow.

The present study provides an insight into understanding the genetic structure of the Balkans. Although the linguistic and cultural diversity found in the region could have acted as an important genetic barrier, Balkan populations have been shown to be genetically homogeneous, and in concordance with the European genetic continuum, using both autosomal and uniparental markers even with the deep levels of resolution conferred by the large set of markers we typed. Linguistic and other cultural differences were probably introduced into genetically homogeneous groups and/or these cultural barriers were not strong enough to prevent genetic flow between populations. However, genetic evidence shows an exception to this pattern: in some particular cases cultural isolation seems to have given rise to small population groups that have become different through drift from the common genetic substrate. This may be the case for some, but not all, Aromun populations.

Acknowledgments

We thank Mònica Vallés, Anna Pérez-Lezaun, Roger Anglada and Stéphanie Plaza, Universitat Pompeu Fabra, Barcelona, for technical support and advice. We especially acknowledge Mary-Claire King, University of Washington; and Vincent Macaulay, University of Glasgow, for details of some published Balkan mtDNA sequences. Funding was provided by project B0S 2001-0794 granted by the Spanish Ministry of Science and Technology.

References

- Achilli, A., Rengo, C., Magri, C., Battaglia, V., Olivieri, A., Scozzari, R., Cruciani, F., Zeviani, M., Briem, E., Carelli, V., Moral, P., Dugoujon, J. M., Roostalu, U., Loogvali, E. L., Kivisild, T., Bandelt, H. J., Richards, M., Villems, R., Santachiara-Benerecetti, A. S., Semino, O. & Torroni, A. (2004) The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. *Am J Hum Genet* **75**, 910–918.
- Bandelt, H. J., Forster, P. & Rohl, A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* **16**, 37–48.
- Barac, L., Pericic, M., Klaric, I. M., Janicijevic, B., Parik, J., Rootsi, S. & Rudan, P. (2003) Y chromosome STRs in Croatians. *Forensic Sci Int* **138**, 127–133.
- Barbarii, L. E., Rolf, B. & Dermengiu, D. (2003) Y-chromosomal STR haplotypes in a Romanian population sample. *Int J Legal Med* **117**, 312–315.
- Belledi, M., Poloni, E. S., Casalotti, R., Conterio, F., Mikerezi, I., Tagliavini, J. & Excoffier, L. (2000) Maternal and paternal lineages in Albania and the genetic structure of Indo-European populations. *Eur J Hum Genet* **8**, 480–486.
- Blanco, P., Shlumukova, M., Sargent, C. A., Jobling, M. A., Affara, N. & Hurles, M. E. (2000) Divergent outcomes of intrachromosomal recombination on the human Y chromosome: male infertility and recurrent polymorphism. *J Med Genet* **37**, 752–758.
- Bosch, E., Lee, A. C., Calafell, F., Arroyo, E., Henneman, P., de Knijff, P. & Jobling, M. A. (2002) High resolution Y chromosome typing: 19 STRs amplified in three multiplex reactions. *Forensic Sci Int* **125**, 42–51.
- Caglia, A., Dobosz, M., Boschi, I., d'Aloja, E. & Pascali, V. L. (1998) Increased forensic efficiency of a STR-based Y-specific haplotype by addition of the highly polymorphic DYS385 locus. *Int J Legal Med* **111**, 142–146.
- Calafell, F., Underhill, P., Tolun, A., Angelicheva, D. & Kalaydjieva, L. (1996) From Asia to Europe: mitochondrial DNA sequence variability in Bulgarians and Turks. *Ann Hum Genet* **60**, 35–49.
- Cavalli-Sforza, L. L., Menozzi, P. & Piazza, A. (1994) *The history and geography of human genes*. Princeton University Press, Princeton, NJ.
- Chikhi, L., Destro-Bisol, G., Bertorelle, G., Pascali, V. & Barbujani, G. (1998) Clines of nuclear DNA markers suggest a largely neolithic ancestry of the European gene pool. *Proc Natl Acad Sci U S A* **95**, 9053–9058.
- Cinnioglu, C., King, R., Kivisild, T., Kalfoglu, E., Atasoy, S., Cavalleri, G. L., Lillie, A. S., Roseman, C. C., Lin, A. A., Prince, K., Oefner, P. J., Shen, P., Semino, O., Cavalli-Sforza, L. L. & Underhill, P. A. (2004) Excavating Y-chromosome haplotype strata in Anatolia. *Hum Genet* **114**, 127–148.
- Comas, D., Calafell, F., Mateu, E., Pérez-Lezaun, A., Bosch, E. & Bertranpetit, J. (1997) Mitochondrial DNA variation and the origin of the Europeans. *Hum Genet* **99**, 443–449.
- Comas, D., Schmid, H., Brauer, S., Flaiz, C., Busquets, A., Calafell, F., Bertranpetit, J., Scheil, H. G., Huckenbeck, W., Efremovska, L. & Schmidt, H. (2004) Alu insertion polymorphisms in the Balkans and the origins of the Aromuns. *Ann Hum Genet* **68**, 120–127.
- Cruciani, F., La Fratta, R., Santolamazza, P., Sellitto, D., Pascone, R., Moral, P., Watson, E., Guida, V., Colomb, E. B., Zaharova, B., Lavinha, J., Vona, G., Aman, R., Cali, F., Akar, N., Richards, M., Torroni, A., Novelletto, A. & Scozzari, R. (2004) Phylogeographic analysis of haplogroup E3b (E-M215) y chromosomes reveals multiple migratory events within and out of Africa. *Am J Hum Genet* **74**, 1014–1022.

- Di Giacomo, F., Luca, F., Anagnou, N., Ciavarella, G., Corbo, R. M., Cresta, M., Cucci, F., Di Stasi, L., Agostiano, V., Giparaki, M., Loutradis, A., Mammi, C., Michalodimitrakis, E. N., Papola, F., Pedicini, G., Plata, E., Terrenato, L., Tofanelli, S., Malaspina, P. & Novelletto, A. (2003) Clinal patterns of human Y chromosomal diversity in continental Italy and Greece are dominated by drift and founder effects. *Mol Phylogenet Evol* **28**, 387–395.
- Excoffier, L., Smouse, P. E. & Quattro, J. M. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479–491.
- Furedi, S., Woller, J., Padar, Z. & Angyal, M. (1999) Y-STR haplotyping in two Hungarian populations. *Int J Legal Med* **113**, 38–42.
- Greenacre, M. (1992) Correspondence analysis in medical research. *Stat Methods Med Res* **1**, 97–117.
- Huckenbeck, W., Scheil, H. G., Schmidt, H. D., Efremovska, L. & Xirotiris, N. (2001) Population genetic studies in the Balkans. II. DNA-STR-systems. *Anthropol Anz* **59**, 213–225.
- Hurles, M. E., Nicholson, J., Bosch, E., Renfrew, C., Sykes, B. C. & Jobling, M. A. (2002) Y chromosomal evidence for the origins of oceanic-speaking peoples. *Genetics* **160**, 289–303.
- Kayser, M., Brauer, S., Weiss, G., Schiefenhovel, W., Underhill, P. A. & Stoneking, M. (2001) Independent histories of human Y chromosomes from Melanesia and Australia. *Am J Hum Genet* **68**, 173–190.
- Kivisild, T., Tolk, H. V., Parik, J., Wang, Y., Papiha, S. S., Bandelt, H. J. & Villems, R. (2002) The emerging limbs and twigs of the East Asian mtDNA tree. *Mol Biol Evol* **19**, 1737–1751.
- Kong, Q. P., Yao, Y. G., Sun, C., Bandelt, H. J., Zhu, C. L. & Zhang, Y. P. (2003) Phylogeny of east Asian mitochondrial DNA lineages inferred from complete sequences. *Am J Hum Genet* **73**, 671–676.
- Macaulay, V., Richards, M., Hickey, E., Vega, E., Cruciani, F., Guida, V., Scozzari, R., Bonne-Tamir, B., Sykes, B. & Torroni, A. (1999) The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. *Am J Hum Genet* **64**, 232–249.
- Malyarchuk, B. A., Grzybowski, T., Derenko, M. V., Czarny, J., Drobic, K. & Miscicka-Sliwka, D. (2003) Mitochondrial DNA variability in Bosnians and Slovenians. *Ann Hum Genet* **67**, 412–425.
- Nasidze, I., Schadlich, H. & Stoneking, M. (2003) Haplotypes from the Caucasus, Turkey and Iran for nine Y-STR loci. *Forensic Sci Int* **137**, 85–93.
- Owens, K. N., Harvey-Blankenship, M. & King, M. C. (2002) Genomic sequencing in the service of human rights. *Int J Epidemiol* **31**, 53–58.
- Paracchini, S., Arredi, B., Chalk, R. & Tyler-Smith, C. (2002) Hierarchical high-throughput SNP genotyping of the human Y chromosome using MALDI-TOF mass spectrometry. *Nucleic Acids Res* **30**, e27.
- Parreira, K. S., Lareu, M. V., Sanchez-Diz, P., Skitsa, I. & Carracedo, A. (2002) DNA typing of short tandem repeat loci on Y-chromosome of Greek population. *Forensic Sci Int* **126**, 261–264.
- Ploski, R., Wozniak, M., Pawlowski, R., Monies, D. M., Branicki, W., Kupiec, T., Kloosterman, A., Dobosz, T., Bosch, E., Nowak, M., Lessig, R., Jobling, M. A., Roewer, L. & Kayser, M. (2002) Homogeneity and distinctiveness of Polish paternal lineages revealed by Y chromosome microsatellite haplotype analysis. *Hum Genet* **110**, 592–600.
- Richards, M., Macaulay, V., Hickey, E., Vega, E., Sykes, B., Guida, V., Rengo, C., Sellitto, D., Cruciani, F., Kivisild, T., Villems, R., Thomas, M., Rychkov, S., Rychkov, O., Rychkov, Y., Golge, M., Dimitrov, D., Hill, E., Bradley, D., Romano, V., Cali, F., Vona, G., Demaine, A., Papiha, S., Triantaphyllidis, C., Stefanescu, G., Hatina, J., Belledi, M., Di Rienzo, A., Novelletto, A., Oppenheim, A., Norby, S., Al-Zaheri, N., Santachiara-Benerecetti, S., Scozzari, R., Torroni, A. & Bandelt, H. J. (2000) Tracing European founder lineages in the Near Eastern mtDNA pool. *Am J Hum Genet* **67**, 1251–1276.
- Richards, M., Macaulay, V., Torroni, A. & Bandelt, H. J. (2002) In search of geographical patterns in European mitochondrial DNA. *Am J Hum Genet* **71**, 1168–1174.
- Robino, C., Varacalli, S., Gino, S., Chatzikiyakidou, A., Kouvatzi, A., Triantaphyllidis, C., Di Gaetano, C., Crobu, F., Matullo, G., Piazza, A. & Torre, C. (2004) Y-chromosomal STR haplotypes in a population sample from continental Greece, and the islands of Crete and Chios. *Forensic Sci Int* **145**, 61–64.
- Rosser, Z. H., Zerjal, T., Hurles, M. E., Adojaan, M., Alavantic, D., Amorim, A., Amos, W., Armenteros, M., Arroyo, E., Barbujani, G., Beckman, G., Beckman, L., Bertranpetit, J., Bosch, E., Bradley, D. G., Brede, G., Cooper, G., Corte-Real, H. B., de Knijff, P., Decorte, R., Dubrova, Y. E., Evgrafov, O., Gilissen, A., Glisic, S., Golge, M., Hill, E. W., Jeziorowska, A., Kalaydjieva, L., Kayser, M., Kivisild, T., Kravchenko, S. A., Krumina, A., Kucinskas, V., Lavinha, J., Livshits, L. A., Malaspina, P., Maria, S., McElreavey, K., Meitinger, T. A., Mikelsaar, A. V., Mitchell, R. J., Nafa, K., Nicholson, J., Norby, S., Pandya, A., Parik, J., Patsalis, P. C., Pereira, L., Peterlin, P., Pielberg, G., Prata, M. J., Previdere, C., Roewer, L., Rootsi, S., Rubinsztein, D. C., Saillard, J., Santos, F. R., Stefanescu, G., Sykes, B. C., Tolun, A., Villems, R., Tyler-Smith, C. & Jobling, M. A. (2000) Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet* **67**, 1526–1543.

- Schmidt, H. S., Scheil, H.-G. & Scheffrahn, W. (2000) The History and Genetics of the Aromun Populations. *Biennial Books of EAA 2000* **1**, 29–37.
- Schmidt, H., Efremskova, L. & Handziski, Z. (2001) Isonymy, consanguinity and repeated pairs of surnames in Aromun populations. *Anthropol Anz* **59**, 193–202.
- Schneider, S. K., Roessli, D. & Excoffier, L. (1996) *Arlequin (ver.1.0): a software environment for the analysis of population genetics data*. University of Geneva, Geneva, Switzerland.
- Semino, O., Magri, C., Benuzzi, G., Lin, A. A., Al-Zahery, N., Battaglia, V., Maccioni, L., Triantaphyllidis, C., Shen, P., Oefner, P. J., Zhivotovsky, L. A., King, R., Torroni, A., Cavalli-Sforza, L. L., Underhill, P. A. & Santachiara-Benerecetti, A. S. (2004) Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the neolithization of Europe and later migratory events in the Mediterranean area. *Am J Hum Genet* **74**, 1023–1034.
- Semino, O., Passarino, G., Oefner, P. J., Lin, A. A., Arbuzova, S., Beckman, L. E., De Benedictis, G., Francalacci, P., Kouvatsi, A., Limborska, S., Marcikiae, M., Mika, A., Mika, B., Primorac, D., Santachiara-Benerecetti, A. S., Cavalli-Sforza, L. L. & Underhill, P. A. (2000) The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. *Science* **290**, 1155–1159.
- Simoni, L., Calafell, F., Pettener, D., Bertranpetit, J. & Barbujani, G. (2000) Geographic patterns of mtDNA diversity in Europe. *Am J Hum Genet* **66**, 262–278.
- Vigilant, L., Pennington, R., Harpending, H., Kocher, T. & Wilson, A. (1989) Mitochondrial DNA sequences in single hairs from a southern African population. *Am J Hum Genet* **86**, 9350–9354.
- Wells, R. S., Yuldashova, N., Ruzibakiev, R., Underhill, P. A., Evseeva, I., Blue-Smith, J., Jin, L., Su, B., Pitchappan, R., Shanmugalakshmi, S., Balakrishnan, K., Read, M., Pearson, N. M., Zerjal, T., Webster, M. T., Zholoshvili, I., Jamarjashvili, E., Gambarov, S., Nikbin, B., Dostiev, A., Aknazarov, O., Zalloua, P., Tsoy, I., Kitaev, M., Mirrakhimov, M., Chariev, A. & Bodmer, W. F. (2001) The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proc Natl Acad Sci U S A* **98**, 10244–10249.
- Whitfield, L. S., Sulston, J. E. & Goodfellow, P. N. (1995) Sequence variation of the human Y chromosome. *Nature* **378**, 379–380.
- Y Chromosome Consortium (2002) A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Res* **12**, 339–348.
- Zaharova, B., Andonova, S., Gilissen, A., Cassiman, J. J., Decorte, R. & Kremensky, I. (2001) Y-chromosomal STR haplotypes in three major population groups in Bulgaria. *Forensic Sci Int* **124**, 182–186.

Received: 22 July 2005

Accepted: 3 October 2005

Appendix II Primers (5' to 3') used in the SNaPshot protocol**Mitochondrial DNA***Amplification primers*

L7008: ACTAGACATCGTACTACACG

H7896: G TACTCGTACCTTCAGTACC

L10373: CCCTAAGTCTGGCCTATGAG

H12744: CGATGAACAGTTGGAATAGG

Typing primers

L7028X: CACGTACTACGTTGTAGC

H10400X: TGTTTAAACTATATACCAATTC

L10873X: T₉CCACAGCCTAATTATTAGCATCATCCCL11251X: T₁₈GGCTCCCTTCCCCTACTCATCGCACTL11719X: T₁₅CAGTCATTCTCATAATCGCCACGG

L12308X: CAGCTATCCATTGGTCTTAGGCCCCAA

L12705X: AACATTAATCAGTTCTTCAAATATCTACTCAT

Y chromosome*Zip code primers*

ZipALg1: GGAGCACGCTATCCCGTTAGAC

ZipBLg2: CGCTGCCAACTACCGCACATG

Amplification primers

M69F: ZipALg1-TGAAGGAATCAGCCATTT CA

M69R: ZipBLg2-TGGCATGAAGCATGTAAGGA

M201F: ZipALg1-TTGTGTGTGTATGCATTTGTTGA

M201R: ZipBLg2-ACATCATGGTGTGACGAACG

P2F: ZipALg1-GAGAATCAGCTCCAGCCATC

P2R: ZipBLg2-TTCTCTCATGAGGGTTTTGGA

M207F: ZipALg1-GGGGCAAATGTAAGTCAAGC

M207R: ZipBLg2-TCCTCTCTGAAATGCCGAAT

PN25F: ZipALg1-AACCTGGAGCATGATCACAC

PN25R: ZipBLg2-GGACCATCACCTGGGTAAAGT

M75F: ZipALg1-GACCAAGAATTTTTTCAGAAAGTGG

M75R: ZipBLg2-GGGTGACAGCGTGAATCTCT

Typing primers

Multiplex I

M89: A₄GGCAAAGTGAGAGATM172: A₇TAATTGAAGACCTTTTAAGTM69: A₁₁GGCTGTTTACACTCCTGAAAM201: A₅TCATCCAACACTAAGTACCTATTACGAAAAM170: A₁₃TGAGACACAACCCACACTGAAAAAAAM9: A₁₃CTGCAAAGAAACGGCCTAAGATGGTTGAAT12f2: A₂₁TCCCTTCCTTACACCTTATACAAAAM145: A₂₈TCCTAGACACCAGAAAGAAAGGC

Multiplex II

M173: A₂TCAAGGGCATTTAGAACM45: A₆GTGAAAAATTATAGATASRY10831: A₉TGAACCTTGAAAATGTTAM207: A₁₁TAAGTCAAGCAAGAAATTTAM17: A₁₄GTGGTTGCTGGTTGTTACGGGPN25: A₁₃TGAATTATCTGCCTGAAACCTGCCTG

Multiplex III

M96: A₃CAGGTCTCTCATAATAP2: A₆TGCCCCCTAGGAGGAGAAM123: A₉CTAGGTATTCAGGCGATGM75: A₁₀GACAATTATCAAACCACATCCM78: A₁₇TGAAATATTTGGAAGGGCM81: A₁₉GTGTGAGTATACTCTATGACM33: A₂₅TCTCATAAGTTACTGTTAM35: A₂₈CGGAGTCTCTGCCTGTGTC

Appendix III Y-STR haplotypes by haplogroup in the Balkans

HG	HT	DYS19	DYS389I	DYS389II	DYS390	DYS391	DYS392	DYS393	DYS385	DYS388	DYS434	DYS435	DYS436	DYS437	DYS438	DYS439	DYS460	DYS461	DYS462	AAA	AAD	AMK	AMS	ARO	GRE	ALB	MAC	ROMP	ROMC	TOTAL		
C	1	15	14	31	24	10	11	13	13-17	13	12	11	11	14	11	14	10	13	11					1					1			
	2	15	14	31	24	10	11	13	13-17	13	12	11	12	14	11	14	10	13	11					1					1			
E1	3	16	12	30	22	10	11	13	14-15	12	10	11	12	16	10	12	9	12	13									1	1			
	4	13	11	27	24	10	11	13	16-19	12	11	11	12	14	10	12	9	12	12		1								1			
E3b1	5	13	12	29	24	10	11	13	16-17	12	11	11	12	14	10	12	9	12	12							1			1			
	6	13	12	29	24	10	11	14	16-18	12	11	11	12	14	10	12	9	12	12						1				1			
	7	13	12	31	23	10	11	13	16-18	12	11	11	12	14	10	11	9	12	12						1				1			
	8	13	13	29	24	9	11	13	16-19	12	11	11	12	14	10	12	9	12	12						1				1			
	9	13	13	29	24	10	11	13	16-18	12	11	11	12	14	10	12	9	12	12			1					1		1			
	10	13	13	29	24	10	11	13	16-19	12	11	11	12	14	10	12	9	12	12		3						1		6			
	11	13	13	30	24	10	10	13	17-18	12	11	11	12	14	10	12	9	12	12									1	1			
	12	13	13	30	24	10	11	13	15-18	12	11	11	12	14	10	12	9	12	12									1	1			
	13	13	13	30	24	10	11	13	15-20	12	11	11	12	14	10	12	9	12	12						1				1	1		
	14	13	13	30	24	10	11	13	16-17	12	11	11	12	14	10	12	9	12	12									1	1	1		
	15	13	13	30	24	10	11	13	16-18	12	11	11	12	14	10	9	9	12	12								1		1	1		
	16	13	13	30	24	10	11	13	16-18	12	11	11	12	14	10	11	9	12	12							1			1	1		
	17	13	13	30	24	10	11	13	16-18	12	11	11	12	14	10	12	9	12	12			1			3	1	1	1	7	1		
	18	13	13	30	24	10	11	13	16-18	12	11	11	12	14	10	12	9	13	12								1			1	1	
	19	13	13	30	24	10	11	13	16-18	12	11	11	12	14	11	11	9	12	12		2						1			1	1	
	20	13	13	30	24	10	11	13	16-18	12	11	11	12	14	12	12	9	11	12						1					2	1	
	21	13	13	30	24	10	11	13	16-19	12	11	11	12	14	10	12	9	11	12			1								1	1	
	22	13	13	30	24	10	11	13	16-19	12	11	11	12	14	10	12	9	12	12			2								1	1	
	23	13	13	30	24	10	11	13	16-19	13	11	11	12	14	10	12	9	12	12							1				2	1	
	24	13	13	30	24	10	11	13	17-17	12	11	11	12	14	10	13	9	12	12								1				1	1
	25	13	13	30	24	10	11	13	17-18	12	11	11	12	14	10	12	9	12	12					1							1	1
	26	13	13	30	24	10	11	14	16-16	12	11	11	12	14	10	12	10	12	12							1					1	1
	27	13	13	30	24	10	11	14	16-18	12	11	11	12	14	10	12	9	13	12							1					1	1
	28	13	13	30	24	11	11	13	16-18	12	11	11	12	14	10	12	9	12	12					6	2						8	1
	29	13	13	30	24	11	11	13	16-18	12	11	11	12	14	10	13	9	12	12				1								1	1
	30	13	13	30	25	10	11	13	15-18	12	11	11	12	14	10	11	9	12	12		1										1	1
	31	13	13	30	25	11	11	13	17-17	12	11	11	12	14	10	11	9	11	12							1					1	1
	32	13	13	31	24	10	11	13	16-16	11	11	11	12	14	10	12	9	13	12								1				1	1
	33	13	13	31	24	10	11	13	17-18	12	10	11	12	14	10	12	9	12	12												1	1
	34	13	13	31	24	11	11	13	16-18	12	11	11	12	14	10	12	9	12	12										1		1	1
	35	13	13	31	25	10	11	13	15-18	12	11	11	12	14	10	11	9	12	12					1							1	1

Appendix III Continued.

HG	HT	DYS19	DYS389I	DYS389II	DYS390	DYS391	DYS392	DYS393	DYS385	DYS388	DYS434	DYS435	DYS436	DYS437	DYS438	DYS439	DYS460	DYS461	DYS462	AAA	AAD	AMK	AMS	ARO	GRE	ALB	MAC	ROMP	ROMC	TOTAL
	36	13	14	30	24	10	11	13	16-19	12	11	11	12	14	10	12	9	12	12				2						2	
	37	13	14	31	24	10	11	13	16-18	12	11	11	12	14	10	12	9	12	12						1				1	
	38	13	14	31	24	10	11	13	17-18	12	11	11	12	14	10	11	9	12	12	2									2	
	39	13	14	31	24	10	12	13	17-20	12	11	11	12	14	10	12	9	12	12						2				2	
	40	14	13	30	24	10	11	13	17-17	12	11	11	12	14	10	11	9	11	12						1				1	
	41	14	13	30	24	10	11	14	17-18	12	11	11	12	14	10	11	11	12	12										1	
	42	14	13	30	24	11	11	13	17-19	12	11	11	12	14	10	11	11	12	12								2		2	
	43	14	13	31	24	10	11	13	16-17	12	11	11	12	14	10	12	9	12	12							1			1	
	44	14	13	31	24	10	11	13	16-18	12	11	11	12	14	10	11	9	12	12										1	
	45	14	13	31	24	10	11	14	17-18	12	11	11	12	14	10	11	11	12	12										1	
E3b2	46	13	13	29	24	9	11	13	14-14	12	11	11	12	14	10	10	11	13	12										1	
E3b3	47	13	13	31	24	9	11	14	17-17	12	11	11	12	14	10	13	10	11	12								1		1	
	48	13	14	32	25	10	11	13	16-16	12	11	11	12	14	11	11	10	11	12								1		1	
	49	13	14	33	24	9	11	13	16-18	12	11	11	12	14	10	11	10	11	12					1					1	
G	50	14	12	30	23	10	11	13	13-14	12	11	11	12	16	10	11	11	12	12								1		1	
	51	15	12	28	22	10	11	14	12-15	12	11	11	12	16	10	12	12	11	12										1	
	52	15	12	28	22	10	11	14	14-14	12	11	11	12	16	11	11	11	11	12										1	
	53	15	12	29	21	11	11	14	14-16	12	11	11	12	16	10	12	9	11	12								1		1	
	54	15	12	29	22	10	10	14	15-17	12	11	11	12	16	10	13	10	12	12									1	1	
	55	15	12	29	22	10	11	14	13-14	13	11	11	12	16	10	11	11	12	12								1		1	
	56	15	12	29	22	11	11	14	14-15	13	11	11	12	16	10	11	10	11	12						1				1	
	57	15	12	29	23	10	11	13	13-14	12	11	11	12	16	10	11	11	13	12								1		1	
	58	15	12	29	23	11	11	13	14-15	12	11	11	12	16	10	11	10	12	12										1	
	59	15	12	31	22	10	11	14	12-15	12	11	11	12	16	10	12	12	11	12										1	
	60	15	13	29	22	10	11	14	13-14	13	11	11	12	16	10	11	10	11	12									1	1	
	61	15	13	30	21	11	11	14	13-16	12	11	11	12	16	10	12	9	11	12									1	1	
	62	15	13	30	23	10	12	13	15-15	12	11	11	12	16	10	11	11	12	11							1			1	
	63	15	14	30	22	10	11	14	13-14	12	10	11	12	16	10	12	11	13	12								1		1	
	64	15	14	30	24	10	12	14	14-14	12	11	11	12	15	10	10	10	10	11										1	
	65	15	14	30	24	10	12	14	14-14	12	11	11	12	15	10	10	10	10	12		1								1	
	66	16	12	29	23	10	11	14	13-14	12	11	11	12	16	10	11	11	11	12									1	3	
	67	15	13	29	22	10	11	12	15-17	12	11	11	12	14	9	11	10	12	11						2				1	
H	68	15	14	30	22	10	11	12	15-17	12	11	11	12	14	9	11	11	12	11								1		2	
	69	14	12	27	23	10	11	13	13-14	14	11	11	12	16	10	11	10	12	12					1					1	
I	70	14	12	28	22	10	11	13	14-14	14	11	10	12	16	10	12	11	12	12										1	

Appendix III Continued.

HT	DVS19	DVS891	DVS3891	DVS390	DVS31	DVS392	DVS393	DVS385	DVS388	DVS434	DVS435	DVS436	DVS437	DVS438	DVS439	DVS460	DVS461	DVS462	AAA	AAD	AMR	AMS	ARO	GRE	ALB	MAC	ROMP	ROMC	TOTAL
71	14	12	28	22	11	11	12	14+14	14	11	11	12	16	10	11	10	12	12										1	1
72	14	12	28	23	10	11	13	13-14	14	11	11	12	16	10	11	10	11	11						1				1	1
73	14	12	28	23	10	11	13	13-14	14	11	11	12	16	10	11	10	12	12		2								4	4
74	14	12	28	23	10	11	13	13-15	14	11	11	12	16	10	11	10	12	12							1			1	1
75	14	12	28	23	10	11	13	14-14	14	11	11	12	16	10	11	10	12	12										1	1
76	14	12	28	23	10	11	14	13-14	14	11	11	12	16	10	10	10	12	12			1							1	1
77	14	12	29	22	10	11	13	14-16	14	11	11	12	16	10	12	10	12	12							1			1	1
78	14	13	29	22	10	11	13	14-15	14	11	11	12	16	10	11	11	12	12								1		1	1
79	15	12	28	22	10	12	13	13-14	14	11	11	12	16	10	12	10	12	12			1							1	1
80	15	12	28	22	10	12	13	14-14	14	11	11	12	16	10	11	10	12	12			1							1	1
81	15	13	29	22	10	13	15	16-16	13	11	11	12	14	10	11	11	12	12								1		1	1
82	15	13	29	23	10	12	14	15-18	13	12	11	12	14	10	11	10	12	12				1						1	1
83	15	13	30	23	10	12	14	14-18	13	12	11	12	14	10	11	10	12	12				1						1	1
84	15	13	30	23	10	12	14	15-18	13	12	11	12	14	10	11	10	12	12				1						1	1
85	15	13	30	23	11	11	13	14-15	13	11	11	12	15	10	14	10	12	12					4					4	12
86	15	13	30	24	11	11	14	14-15	13	11	11	12	15	10	12	10	12	12			3							3	3
87	15	13	30	24	12	11	13	14-15	13	11	11	12	15	10	13	10	12	12										1	1
88	15	13	31	24	10	11	13	14-15	13	11	11	12	15	10	13	10	12	11			1							1	1
89	15	13	31	24	10	11	14	14-15	13	11	11	12	15	10	13	10	12	11			1							1	1
90	15	13	31	24	11	11	13	14-15	13	11	11	12	15	10	12	10	12	12				1						1	1
91	15	13	31	24	11	11	13	14-16	13	11	11	12	15	10	12	10	12	12										1	1
92	15	13	32	24	11	11	13	14-15	13	11	11	12	15	10	13	10	12	12			3							3	6
93	15	14	31	23	10	12	14	14-16	13	11	11	12	14	10	12	11	12	12										2	2
94	15	14	33	24	11	11	13	14-15	13	11	11	12	14	10	12	10	12	12										1	1
95	16	13	29	23	10	12	14	15-15	13	12	11	12	15	8	12	11	12	12									1	1	1
96	16	13	29	23	10	12	14	15-15	13	12	11	12	15	10	12	11	12	12									1	1	1
97	16	13	29	23	11	12	14	15-15	13	11	11	12	15	10	11	11	12	12									1	1	1
98	16	13	30	24	11	11	13	14-15	13	11	11	12	15	10	13	10	12	12									1	4	4
99	16	13	30	24	11	11	13	14-15	13	11	11	12	15	10	14	10	12	12							1			1	1
100	16	13	30	24	11	11	13	15-15	13	11	11	12	15	10	14	10	12	12								1		1	1
101	16	13	31	23	11	11	13	14-16	13	11	11	12	15	10	13	10	11	12								1		1	1
102	16	13	31	24	10	11	13	14-15	13	11	11	12	15	10	12	10	12	12								1		1	1
103	16	13	31	24	11	11	13	13-14	12	11	11	12	15	10	13	11	12	12										1	1
104	16	13	31	24	11	11	13	14-15	13	11	11	12	14	10	12	10	12	12									1	1	1
105	16	13	31	24	11	11	13	14-15	13	11	11	12	15	10	12	11	12	12									1	1	1

Appendix III Continued.

HG	HT	DYS19	DYS389I	DYS389II	DYS390	DYS391	DYS392	DYS393	DYS385	DYS388	DYS434	DYS435	DYS436	DYS437	DYS438	DYS439	DYS460	DYS461	DYS462	AAA	AAD	AMK	AMS	ARO	GRE	ALB	MAC	ROMP	ROMC	TOTAL
	106	16	13	31	24	11	11	13	14-15	13	11	11	12	15	10	13	10	12	12	1		1							4	
	107	16	13	31	24	11	11	13	14-15	13	11	11	12	15	10	14	10	12	12			1							1	
	108	16	13	31	24	11	11	13	14-15	13	11	11	12	15	12	12	10	12	12					1					1	
	109	16	13	31	24	11	11	13	14-16	13	11	11	12	14	10	13	10	12	12						1				1	
	110	16	13	31	24	11	11	13	14-16	13	11	11	12	15	10	13	10	12	12						2	1			3	
	111	16	13	31	24	11	11	14	15-15	13	11	11	12	15	10	13	10	12	12						1				1	
	112	16	13	31	24	12	11	13	14-15	13	11	11	12	15	10	15	10	12	12					1					1	
	113	16	13	31	25	10	11	13	14-15	13	11	11	12	15	10	13	10	12	12					1					1	
	114	16	13	31	25	11	11	13	15-15	13	11	11	12	15	10	12	10	12	12							1			1	
	115	16	13	32	24	10	11	13	14-15	13	11	11	12	15	10	12	10	12	12					1					1	
	116	16	13	32	24	11	11	12	14-15	13	11	11	12	15	10	13	10	12	11								1		1	
	117	16	13	32	24	11	11	13	14-15	13	11	11	12	15	10	12	10	12	12							1			2	
	118	16	13	32	24	11	11	13	14-15	13	12	11	12	15	10	12	10	12	12					1					1	
	119	16	13	32	24	12	11	13	14-15	13	11	11	12	15	10	14	10	12	12								1		1	
	120	16	13	33	24	11	11	13	14-15	13	11	11	12	15	10	14	10	13	12									1	1	
	121	16	13	34	24	11	11	13	14-15	13	11	11	12	15	10	14	10	12	12					1					1	
	122	16	14	33	24	11	11	13	14-15	13	11	11	12	14	10	13	10	12	12							1			1	
	123	16	14	33	24	11	11	13	14-15	13	11	11	12	15	10	13	10	12	12							1			1	
	124	17	13	30	24	10	11	13	14-15	13	11	11	12	15	9	14	10	12	12								1		1	
	125	17	13	30	24	10	11	13	14-15	13	11	11	12	15	10	13	10	12	12				1						1	
	126	17	13	30	24	10	11	13	14-16	13	11	11	12	15	10	13	11	12	15										1	
	127	17	13	30	24	11	11	13	14-15	13	12	11	12	15	10	13	10	12	12					1					1	
	128	17	13	31	24	10	11	13	14-15	13	11	11	12	15	10	13	10	12	12										1	
	129	17	13	31	24	11	11	12	14-15	13	11	11	12	15	10	12	10	12	12							1			1	
	130	17	13	31	24	11	11	13	14-15	13	11	11	12	15	10	12	10	12	12								1		1	
	131	17	13	31	24	11	11	13	14-15	13	11	11	12	15	10	13	10	12	12							1			2	
	132	17	13	31	24	11	11	13	14-15	13	11	11	12	16	10	13	10	12	12									1	1	
	133	17	13	31	24	11	11	14	14-16	13	11	11	12	15	10	14	10	12	12									1	1	
	134	17	13	32	22	11	11	13	14-16	13	11	11	12	15	10	13	10	13	12					1					1	
	135	17	13	32	24	10	11	13	14-15	13	11	11	12	15	10	13	10	12	12									1	1	
	136	14	13	28	23	10	11	12	12-18	13	11	11	11	14	10	11	10	12	13									1	1	
	137	14	13	29	25	10	11	13	12-18	16	11	11	12	14	10	11	10	11	11								1		1	
	138	15	13	30	23	10	11	12	13-18	16	11	11	12	14	10	11	11	12	11						1				1	
	139	14	11	27	23	11	11	12	13-17	15	11	11	12	15	9	11	10	13	11									1	1	
	140	14	12	28	23	10	11	12	13-17	14	11	11	12	15	9	11	9	13	11									1	1	

J[8]2

J2

Appendix III Continued.

HQ	HT	DVS19	DVS891	DVS891	DVS390	DVS391	DVS92	DVS93	DVS85	DVS88	DVS44	DVS45	DVS46	DVS47	DVS48	DVS49	DVS60	DVS61	DVS62	AAA	AAD	AMK	AMS	ARO	GRE	ALB	MAC	ROMP	ROMC	TOTAL
141	14	14	12	28	24	10	11	12	14-14	15	11	11	12	16	9	12	11	10	11	11					1				1	
142	14	14	12	28	24	10	11	12	14-17	15	11	11	12	16	9	12	10	10	11	11				1					1	
143	14	14	12	28	24	10	11	12	14-18	15	11	11	10	16	10	13	11	10	11	11					1				1	
144	14	14	12	29	24	10	11	12	14-17	17	11	11	12	16	9	12	11	10	11	11			3						3	
145	14	14	12	29	24	10	11	12	14-17	17	11	11	12	16	9	13	11	10	11	11			2						19	
146	14	14	12	29	24	10	11	12	14-17	17	11	11	12	16	9	13	11	11	11	11			1						1	
147	14	14	12	29	24	10	11	12	14-17	17	11	11	12	16	9	14	11	10	11	11				1					1	
148	14	14	12	29	24	10	11	12	14-17	17	11	11	12	17	9	13	11	10	11	11				1					1	
149	14	14	12	29	24	10	11	12	14-18	17	11	11	12	16	9	12	11	10	11	11			1						1	
150	14	14	12	29	24	11	13	12	14-17	17	11	11	12	16	9	12	11	10	11	11			1						1	
151	14	14	13	29	22	9	11	12	13-17	15	11	11	12	15	9	11	10	12	11	11						1			1	
152	14	14	13	29	23	10	11	12	13-16	15	11	11	12	15	9	11	10	13	11	11				1					1	
153	14	14	13	29	24	10	11	12	13-16	16	11	11	12	14	9	10	10	12	11	11				1					1	
154	14	14	13	29	25	10	11	12	14-15	14	11	11	12	15	10	12	10	12	11	11					1				1	
155	14	14	13	30	22	10	11	12	12-15	17	11	11	12	14	9	11	10	13	11	11								1	1	
156	14	14	13	30	23	10	11	12	13-18	15	11	11	12	15	9	11	11	13	11	11				1					1	
157	14	14	13	30	23	10	11	12	14-15	15	11	11	12	15	9	11	9	11	11	11				1					1	
158	14	14	13	30	23	10	11	13	13-17	15	11	11	12	14	9	11	10	12	11	11						1			1	
159	15	12	27	24	10	11	12	12	14-17	15	11	11	12	16	9	11	10	10	11	11				1					1	
160	15	12	28	24	10	11	12	12	13-17	15	11	11	12	16	9	12	10	10	11	11						1			1	
161	15	12	28	24	10	11	12	12	14-16	15	11	11	12	16	9	12	10	10	11	11						1			1	
162	15	12	28	24	10	11	12	12	14-17	15	11	11	12	16	9	12	11	10	11	11						1			1	
163	15	12	28	24	10	11	12	12	14-17	16	11	11	12	16	9	11	11	10	11	11						1			1	
164	15	12	28	24	10	11	12	12	15-18	15	11	11	12	16	9	12	11	10	11	11				2					2	
165	15	12	28	24	11	11	12	12	17-18	12	11	11	12	14	9	12	11	10	11	11						1			1	
166	15	13	29	23	9	11	12	12	13-16	16	11	11	12	14	9	11	10	13	10	11				2	12	2	1		17	
167	15	13	29	23	10	13	12	12	13-17	15	11	11	12	15	9	12	10	12	11	11			1						1	
168	15	13	29	24	9	11	12	12	13-16	16	11	11	12	14	9	11	10	13	10	11									1	
169	15	13	29	24	10	11	12	12	13-17	16	11	11	12	16	9	12	11	10	11	11						1			1	
170	15	13	29	25	10	11	12	12	13-16	15	11	11	12	15	9	11	11	12	11	11				1					1	
171	15	13	30	23	9	11	12	12	13-16	16	11	11	12	14	9	11	10	13	10	11						1			1	
172	15	13	31	22	10	11	13	13	13-15	15	11	11	12	14	9	12	10	13	11	11							1		1	
173	15	14	30	23	10	11	12	12	13-16	16	11	11	12	14	9	12	10	13	8	11							1		1	
174	16	12	28	24	10	11	12	12	14-16	15	11	11	12	16	9	12	11	10	11	11							1		1	
175	16	13	28	24	10	11	12	12	16-18	15	11	11	12	14	9	12	11	10	13	11							1		1	

Appendix III Continued.

HG	HT	DYS19	DYS389I	DYS389II	DYS390	DYS391	DYS392	DYS393	DYS385	DYS388	DYS434	DYS435	DYS436	DYS437	DYS438	DYS439	DYS460	DYS461	DYS462	AAA	AAD	AMK	AMS	ARO	GRE	ALB	MAC	ROMP	ROMC	TOTAL
	176	16	13	30	23	10	11	12	12-15	15	11	11	12	15	9	11	10	13	11								1		1	
	177	17	13	29	23	9	11	12	13-16	16	11	11	12	14	9	11	10	14	10							1			1	
K(3P)	178	13	13	29	23	10	14	13	14-17	12	11	11	12	15	9	11	11	11	12						1				1	
	179	14	12	30	23	9	14	12	14-16	12	11	11	12	14	9	12	10	11	12							1			1	
	180	14	13	29	23	10	13	13	14-15	12	11	11	12	14	9	10	11	11	13							1			1	
	181	14	13	30	23	9	14	12	14-16	12	11	11	12	14	9	12	10	11	12							1			1	
	182	14	13	30	23	10	13	13	15-16	12	11	11	12	14	9	11	10	11	12						1				1	
	183	15	13	29	23	12	14	14	11-14	12	10	11	12	14	10	10	11	12	12										2	
	184	15	13	29	24	10	13	13	13-16	12	11	11	12	14	9	12	10	11	13										1	
R1*	185	14	12	28	24	11	13	13	11-14	12	11	11	12	15	12	12	11	12	11								1		1	
R1a1	186	15	12	29	25	11	11	14	11-16	12	11	11	12	14	9	11	10	11	11										1	
	187	15	13	29	26	11	11	13	11-14	11	11	11	12	14	11	10	11	11	11										1	
	188	15	13	30	25	11	11	13	11-14	12	11	11	12	14	10	10	11	11	11										1	
	189	15	13	31	25	10	11	13	11-15	12	11	11	12	14	11	10	10	11	11										1	
	190	15	13	31	25	11	11	13	11-14	12	11	11	12	14	11	10	10	12	11							2			2	
	191	15	13	32	26	10	11	13	11-14	12	11	11	12	14	11	10	11	12	11										1	
	192	15	14	31	25	11	11	13	11-14	12	11	11	12	14	11	10	10	11	11										2	
	193	16	12	28	25	10	11	13	11-14	12	11	11	12	14	11	11	10	11	11							1			1	
	194	16	12	30	25	11	11	12	12-14	12	11	11	12	14	11	11	11	11	11										1	
	195	16	13	29	24	10	11	13	10-14	12	11	11	12	14	11	11	11	11	11										1	
	196	16	13	29	24	11	11	13	11-14	12	11	11	12	14	11	11	11	11	11							1			1	
	197	16	13	29	25	10	11	13	11-14	12	11	11	12	14	11	11	11	11	11							1			1	
	198	16	13	29	25	11	11	13	11-13	12	11	11	12	14	11	11	11	12	11										2	
	199	16	13	29	25	11	11	13	11-14	12	11	11	12	14	11	10	11	11	11								1		1	
	200	16	13	30	25	10	11	13	11-14	12	11	11	12	14	11	10	10	11	12								2		2	
	201	16	13	30	25	10	11	13	11-15	12	11	11	12	14	11	12	11	11	11										1	
	202	16	13	30	25	11	11	13	11-13	12	11	11	12	14	11	11	11	13	11										1	
	203	16	13	30	25	11	11	13	11-14	12	11	11	12	14	11	11	11	11	11										1	
	204	16	13	30	25	11	11	13	11-15	12	11	11	12	14	11	10	10	11	11										2	
	205	16	13	30	25	11	11	13	11-16	12	11	11	12	14	11	10	11	11	11									1	1	
	206	16	14	30	25	10	11	13	11-11	12	11	11	12	14	11	11	11	11	11								1		1	
	207	16	14	30	25	10	11	13	11-14	12	11	11	12	14	11	11	11	11	11								1		1	
	208	16	14	31	25	10	11	13	11-14	12	11	11	12	14	11	11	11	11	11										1	
	209	16	14	31	25	11	11	13	11-15	12	11	11	12	14	11	10	11	11	11										1	
	210	17	13	29	25	11	11	13	12-14	12	11	11	12	14	11	10	12	11	11										7	

Appendix III Continued.

HQ	HT	DVS19	DVS3891	DVS3891I	DVS390	DVS391	DVS392	DVS393	DVS385	DVS388	DVS434	DVS435	DVS436	DVS437	DVS438	DVS439	DVS460	DVS461	DVS462	AAA	AAD	AMR	AMS	ARO	GRE	ALB	MAC	ROMP	ROMC	TOTAL
	211	17	13	29	25	12	11	13	12-14	12	11	11	12	14	11	10	12	11	11											1
	212	17	13	29	26	11	11	13	12-14	12	11	11	12	14	11	10	12	11	11											2
	213	17	13	30	24	10	11	13	10-14	12	11	11	12	14	11	10	11	11	11											1
	214	17	13	30	25	10	11	13	10-14	12	11	11	12	14	11	10	11	11	11											1
	215	17	13	30	25	11	11	13	11-13	12	11	11	12	14	11	11	11	12	11											1
	216	17	14	30	25	11	11	13	12-14	12	11	11	12	14	11	11	12	11	11											1
	217	17	14	30	26	11	11	13	12-14	12	11	11	12	14	11	11	12	11	11											1
	218	17	14	31	25	11	11	13	11-14	12	11	11	12	14	11	10	11	11	11											1
	219	17	14	31	26	11	11	13	11-14	12	11	11	12	14	11	10	11	11	11											1
	220	13	13	30	24	11	13	13	11-14	12	11	11	12	15	12	12	11	12	11											1
	221	14	12	28	24	11	13	13	11-14	12	11	11	12	15	12	13	11	12	11											1
	222	14	12	28	26	10	14	12	11-15	12	11	11	12	16	12	13	10	11	11											1
	223	14	13	28	25	10	14	12	11-14	12	11	11	13	15	12	12	11	11	11											1
	224	14	13	29	23	10	13	12	10-14	12	11	11	12	15	12	13	11	11	11											1
	225	14	13	29	23	10	13	13	11-14	12	11	11	12	15	12	12	11	12	11											1
	226	14	13	29	23	11	13	13	11-14	12	11	11	12	15	12	13	11	13	11											1
	227	14	13	29	23	11	15	13	11-14	12	11	11	12	15	12	12	11	12	11											1
	228	14	13	29	24	10	13	12	11-13	12	11	11	12	15	12	13	10	12	11											1
	229	14	13	29	24	10	13	12	11-14	13	11	11	12	15	12	11	11	12	11											1
	230	14	13	29	24	10	13	12	11-14	13	11	11	12	15	12	12	11	12	11											1
	231	14	13	29	24	11	11	12	11-11	12	11	11	12	15	12	12	11	11	11											6
	232	14	13	29	24	11	11	12	11-11	12	11	11	12	15	12	12	12	11	11											2
	233	14	13	29	24	11	13	12	11-13	12	11	11	12	15	12	12	10	12	11											1
	234	14	13	29	24	11	13	12	11-14	12	11	11	12	15	12	12	11	11	11											10
	235	14	13	29	24	11	13	12	11-14	12	11	11	12	15	12	13	11	11	11											1
	236	14	13	29	24	11	13	13	11-11	12	11	11	12	15	12	11	11	11	11											1
	237	14	13	29	24	11	13	13	11-11	12	11	11	12	15	12	12	12	11	11											3
	238	14	13	29	24	11	13	13	11-14	12	11	11	12	15	12	12	11	12	11											1
	239	14	13	29	24	11	13	14	11-11	12	11	11	12	15	12	12	11	11	11											1
	240	14	13	29	25	10	13	12	11-14	12	11	11	12	15	12	12	11	11	11											1
	241	14	13	29	25	11	13	12	11-14	12	11	11	12	15	12	15	10	11	11											1
	242	14	13	29	25	11	13	13	11-11	12	11	11	12	15	12	12	12	11	11											1
	243	14	13	29	26	11	13	13	11-15	12	11	11	12	16	12	13	11	11	11											1
	244	14	13	30	23	10	13	12	11-15	12	11	12	12	15	12	12	11	11	11											1
	245	14	13	30	24	10	11	12	11-11	12	11	11	12	15	12	12	12	11	11											1

R4b

Appendix III Continued.

HG	HT	DYS19	DYS389I	DYS389II	DYS390	DYS391	DYS392	DYS393	DYS385	DYS388	DYS434	DYS435	DYS436	DYS437	DYS438	DYS439	DYS460	DYS461	DYS462	AAA	AAD	AMK	AMS	ARO	GRE	ALB	MAC	ROMP	ROMC	TOTAL
246	14	13	30	24	11	13	12	11-14	12	11	11	12	15	12	12	11	11	11	11					1					1	
247	14	13	30	25	11	14	13	11-14	13	11	11	12	15	12	13	12	13	11	11							1			1	
248	14	13	31	24	11	13	12	11-11	12	12	11	11	12	15	12	12	11	11	11							1			1	
249	14	13	31	24	11	13	13	11-14	12	11	11	11	12	15	12	12	11	12	11	1									1	
250	14	14	30	24	11	13	13	11-13	12	11	11	11	12	15	9	12	11	12	11								1		1	
251	15	13	29	24	11	13	13	12-16	14	11	12	12	15	12	12	12	11	12	11					1					1	
252	15	13	29	25	11	13	12	12-13	12	11	11	11	11	14	12	12	9	12	11					1					1	
253	15	13	29	25	11	13	12	12-13	12	11	11	11	11	14	12	12	10	12	11	1			2	4					7	
254	15	13	29	26	11	13	12	12-13	12	11	11	11	11	14	12	12	10	12	11				1						1	
255	15	13	30	24	10	13	13	13-16	12	11	11	11	12	15	13	12	11	12	11				1						2	
256	15	13	30	24	10	13	13	13-16	12	11	11	11	12	15	13	13	11	12	11				1						1	
257	15	13	30	24	11	13	13	11-16	12	11	11	11	12	15	13	12	11	12	11									1	1	
258	15	13	30	25	11	13	12	12-13	12	11	11	11	11	14	12	12	10	12	11				1						1	
259	15	13	30	25	11	13	12	12-13	12	11	11	11	11	14	12	13	10	13	11					1					1	
260	15	13	31	24	10	13	13	11-16	12	11	11	11	12	15	13	12	11	12	11										1	
261	15	13	31	25	11	13	12	12-13	12	11	11	11	11	14	12	12	10	12	11										1	
262	15	14	30	24	12	13	13	11-14	12	11	11	11	12	15	12	12	10	12	11						1				1	
263	15	14	30	25	11	13	12	10-15	12	11	11	11	12	15	12	12	10	11	11								1		1	